

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 18, 2002, 10:48:18 ; Search time 32.51 Seconds
(without alignments)
1219.728 Million cell updates/sec

Title: US-09-728-309-3

Perfect score: 1855

Sequence: 1 MLKLVIIENMAEIMFLSDL.....YLSPPKMLKFNRCGSKTKK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- A_Geneseq_032802.*
- 1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
 - 2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
 - 3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
 - 4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
 - 5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
 - 6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
 - 7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
 - 8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
 - 9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
 - 10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
 - 11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
 - 12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
 - 13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
 - 14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
 - 15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
 - 16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
 - 17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
 - 18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
 - 19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
 - 20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
 - 21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
 - 22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Length	ID	Description
1	365.5	19.7	310	AAW19107	Rat pheromone rece
2	325.5	17.5	311	AAW19104	Rat pheromone rece
3	313.5	16.9	315	AAW19103	Rat pheromone rece
4	311.5	16.8	311	AAW19109	Rat pheromone rece
5	288.5	15.6	307	AAW19108	Rat pheromone rece
6	286.5	15.4	278	AAW19106	Rat pheromone rece
7	263.5	14.2	310	AAW19105	Rat pheromone rece
8	246	13.3	723	ABG11559	Novel human diapo
9	198.5	10.7	254	AAE06763	Human G-protein co
10	193.5	10.4	394	AAE97828	Human G-protein co
11	192.5	10.4	353	AAE64295	Human GTP-binding

12	192.5	10.4	353	22	AAE20349	Human vomeronasal-
13	192	10.4	313	22	AAE20350	Human vomeronasal-
14	192	10.4	313	22	AAE20351	Human vomeronasal-
15	187	10.1	313	22	AAE20348	Human vomeronasal-
16	187	10.1	313	22	AAE20352	Human vomeronasal-
17	186	10.0	155	22	AAU69499	Human purified sec
18	162.5	8.8	286	22	AAE20353	Chimpanzee vomeron
19	121.5	6.5	69	22	ABBA1363	Peptide #8869 enco
20	121.5	6.5	69	22	ABE25300	Protein #7299 enco
21	121.5	6.5	69	22	AAE62231	Human brain expres
22	121.5	6.5	69	22	AAW75035	Human bone marrow
23	121.5	6.5	69	22	AAE35151	Peptide #9188 enco
24	120.5	6.5	241	22	ABG11558	Novel human diapo
25	119	6.4	121	22	ABE28246	Human peptide #897
26	119	6.4	121	22	ABE33421	Peptide #927 enco
27	119	6.4	121	22	ABE18880	Protein #879 enco
28	119	6.4	121	22	AAE54206	Human brain expres
29	119	6.4	121	22	AAE66600	Human bone marrow
30	119	6.4	121	22	AAE14473	Peptide #907 enco
31	119	6.4	121	22	AAE26886	Peptide #923 enco
32	119	6.4	121	22	AAE02200	Peptide #882 enco
33	119	6.4	124	22	ABE22649	Protein #4648 enco
34	109	5.9	376	17	AAE89359	Frog bombesin-like
35	109	5.9	376	22	AAU07635	Frog bombesin-like
36	101.5	5.5	313	22	AAE87748	Human T2R18 amino
37	100	5.4	254	22	AAE71755	Human olfactory re
38	99.5	5.4	313	22	AAU24663	Human olfactory re
39	99.5	5.4	313	22	AAE71707	Human olfactory re
40	99.5	5.4	313	22	AAE72850	Human olfactory re
41	98	5.3	662	22	AAE01079	CYE 82 protein seq
42	97.5	5.3	321	22	AAE71527	Human olfactory re
43	97.5	5.3	325	22	ABE44531	Human GPCR6a polyp
44	97.5	5.3	325	22	AAU24570	Human olfactory re
45	97	5.2	309	22	AAE87738	Human T2R08 amino

ALIGNMENTS

RESULT 1

AAW19107	AAW19107 standard; Protein; 310 AA.
XX	AC AAW19107;
XX	DT 26-AUG-1997 (first entry)
XX	DE Rat pheromone receptor VN6.
XX	KW Pheromone receptor; vomeronasal sensory neuron; social behaviour;
XX	KW maternal behaviour; reproductive behaviour; fertility;
XX	KW hormone secretion.
XX	OS Rattus sp.
XX	Key Location/Qualifiers
FT	Domain 16..37
FT	/label= I
FT	/note= "predicted position of transmembrane domain
FT	I"
FT	Domain 50..71
FT	/label= II
FT	/note= "predicted position of transmembrane domain
FT	II"
FT	Domain 97..116
FT	/label= III
FT	/note= "predicted position of transmembrane domain
FT	III"
FT	Domain 135..156
FT	/label= IV
FT	/note= "predicted position of transmembrane domain
FT	IV"
FT	Domain 194..212

CC Rat pheromone receptors VN1, VN3, VN4, VN5, VN6, VN7 and VN2
 CC (AAW19103-09) are members of a novel family of presumed
 CC 7-transmembrane domain receptors that are evolutionary independent
 CC of the odorant receptors of the main olfactory epithelium (MOE).
 CC Their amino acid sequences were deduced from cDNA clones (see also
 CC AAT69545-50, VN2 sequence not provided) obtd. by differential
 CC cloning, PCR and hybridisation from single vomeronasal sensory
 CC neurons (VSN). Individual neurons express different complements of
 CC the receptors. VN polypeptides can be expressed in host-vector
 CC systems for use in identifying modulators for control of maternal,
 CC reproductive and social behavior, to increase fertility, control
 CC hormone secretion and to regulate food uptake in humans and
 CC animals.
 XX
 SQ Sequence 311 AA;
 Query Match 17.5%; Score 325.5; DB 18; Length 311;
 Best Local Similarity 29.2%; Pred. No. 2e-27;
 Matches 94; Conservative 49; Mismatches 118; Indels 61; Gaps 8;
 QY 44 TIQIFVPOASFGISANTILLFHIFTFVSHRSKSIDMIISHLSLHILLFTQAILVS 103
 Db 13 tirntfsteigilansflllhfikfirqrsrtdipigllslhllmlmga-fia 71
 QY 104 LDFFGSNTODDLRYKVVIVELNKMVRLSICTPCLLSVLQAI-ISPSIFSLAKLHPSAS 162
 Db 72 idifiswrgwddickflvyisfrglstctcmisvlqaitlsprssclakfkhsph 131
 QY 163 HILGFFLEFSLVNMFGVFCCTLRPPVKGQSSVCHTALFLFAHELHPQEVFTNDF 222
 Db 132 hvscalislsilymfisshllvsinatp-----nltnnf 166
 QY 223 -----EGCHLYRVHGPLRLRGDYF--IQTIR--GYLSAFTQAPAC----- 258
 Db 167 mqvqtscyil----pyslmqsmfstlailrdslslsmvstcymevllchrnqihl 222
 QY 259 -----PRVSPVRASQAIIILVSVFTYVWDTFSESGGVTVINDSLVWLQVIVANSY 312
 Db 223 qgtnlspkaspedratqtilmlmtffvlmsifdsivscsrmylndptsyyiqfgvdy 282
 QY 313 AATISPLMLIYADNQIFKTLQML 334
 Db 283 atvspfvmstekhivnflksm 304
 RESULT 3
 AAW19103
 ID AAW19103 standard; Protein; 315 AA.
 AC AAW19103;
 XX
 DT 26-AUG-1997 (first entry)
 XX
 DE Rat pheromone receptor VN1.
 XX
 KW Pheromone receptor; vomeronasal sensory neuron; social behaviour;
 KW maternal behaviour; reproductive sensory neuron; fertility;
 KW hormone secretion.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT Domain 17..38
 FT /label= I
 FT /note= "predicted position of transmembrane domain
 FT I"
 FT Domain 51..72
 FT /label= II
 FT /note= "predicted position of transmembrane domain
 FT II"
 FT Domain 98..117
 FT /label= III

FT /note= "predicted position of transmembrane domain
 FT III"
 FT Domain 136..157
 FT /label= IV
 FT /note= "predicted position of transmembrane domain
 FT IV"
 FT Domain 195..213
 FT /label= V
 FT /note= "predicted position of transmembrane domain
 FT V"
 FT Domain 241..259
 FT /label= VI
 FT /note= "predicted position of transmembrane domain
 FT VI"
 FT Domain 271..293
 FT /label= VII
 FT /note= "predicted position of transmembrane domain
 FT VII"
 XX W09714790-A1.
 XX 24-APR-1997.
 XX 18-OCT-1996; 36WO-US16637.
 XX 19-OCT-1995; 95US-0005698.
 XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX Axel R, Dulac C;
 XX WPI; 1997-245107/22.
 DR N-PSDB; AAT69545.
 XX Nucleic acid molecule encoding vertebrate pheromone receptor -
 PT useful to identify modulators for control of reproductive and social
 PT behaviour, fertility and hormone secretion
 XX Disclosure; Page 74-75; 123pp; English.
 XX Rat pheromone receptors VN1, VN3, VN4, VN5, VN6, VN7 and VN2
 CC (AAW19103-09) are members of a novel family of presumed
 CC 7-transmembrane domain receptors that are evolutionary independent
 CC of the odorant receptors of the main olfactory epithelium (MOE).
 CC Their amino acid sequences were deduced from cDNA clones (see also
 CC AAT69545-50, VN2 sequence not provided) obtd. by differential
 CC cloning, PCR and hybridisation from single vomeronasal sensory
 CC neurons (VSN). VN1 is expressed only in about 4% of VSN and never
 CC in the MOE. Individual neurons express different complements of
 CC the receptors. VN polypeptides can be expressed in host-vector
 CC systems for use in identifying modulators for control of maternal,
 CC reproductive and social behavior, to increase fertility, control
 CC hormone secretion and to regulate food uptake in humans and
 CC animals.
 XX
 SQ Sequence 315 AA;
 Query Match 16.9%; Score 313.5; DB 18; Length 315;
 Best Local Similarity 29.5%; Pred. No. 4.4e-26;
 Matches 96; Conservative 48; Mismatches 110; Indels 71; Gaps 8;
 QY 49 FYPQASFGISANTILLFHIFTFVSHRSKSIDMIISHLSLHILLFTQAILVSLDFPG 108
 Db 18 ffaeigigvansllllfnfkicgqrsrtdipigllslhllmlmga-fiatdtf1 76
 QY 109 SONTQDLDRLRYKVVIVELNKMVRLSICTPCLLSVLQAI-ISPSIFSLAKLHPSASHILGP 167
 Db 77 swrgwddicklsillyrtfrglsctclslsvlqailspssclakfkhsphhsca 136
 QY 168 FLFSWVLMNMFIVGIFCCTLRPLPVKRGQSSVCHTALFLFAHELHPQEVFTNDF----E 223
 Db 137 ilslsvlymfis-----shllvsila---tpnit---tndfthvtq 171

FT	Domain	I"	
FT		47..68	
FT		/label= II	
FT		/note= "predicted position of transmembrane domain	
FT		II"	
FT	Domain	94..113	
FT		/label= III	
FT		/note= "predicted position of transmembrane domain	
FT		III"	
FT	Domain	132..153	
FT		/label= IV	
FT		/note= "predicted position of transmembrane domain	
FT		IV"	
FT	Domain	191..209	
FT		/label= V	
FT		/note= "predicted position of transmembrane domain	
FT		V"	
FT	Domain	237..255	
FT		/label= VI	
FT		/note= "predicted position of transmembrane domain	
FT		VI"	
FT	Domain	267..289	
FT		/label= VII	
FT		/note= "predicted position of transmembrane domain	
FT		VII"	
XX			
PN	WO9714790-A1.		
XX			
XX	24-APR-1997.		
XX			
XX	18-OCT-1996;	96WO-US16637.	
PF			
XX	19-OCT-1995;	95US-0005698.	
PR			
XX			
PA	{UYCO } UNIV COLUMBIA NEW YORK.		
PI	Axel R, Dulac C;		
XX			
DR	WPI; 1997-245107/22.		
DR	N-PSDB; AAT69550.		
XX			
PT	Nucleic acid molecule encoding vertebrate pheromone receptor -		
PT	useful to identify modulators for control of reproductive and social		
PT	behaviour, fertility and hormone secretion		
XX			
PS	Disclosure; Page 84-86; 123pp; English.		
XX			
CC	Rat pheromone receptors VN1, VN3, VN4, VN5, VN6, VN7 and VN2		
CC	(AAW19103-09) are members of a novel family of presumed		
CC	7-transmembrane domain receptors that are evolutionary independent		
CC	of the odorant receptors of the main olfactory epithelium (MOE).		
CC	Their amino acid sequences were deduced from cDNA clones (see also		
CC	AAT69545-50, VN2 sequence not provided) obtd. by differential		
CC	cloning, PCR and hybridisation from single vomeronasal sensory		
CC	neurons (VSN). Individual neurons express different complements of		
CC	the receptors. VN polypeptides can be expressed in host-vector		
CC	systems for use in identifying modulators for control of maternal,		
CC	reproductive and social behavior, to increase fertility, control		
CC	hormone secretion and to regulate food uptake in humans and		
CC	animals.		
XX			
SQ	Sequence 307 AA;		
	Query Match 15.6%; Score 288.5; DB 18; Length 307;		
	Best Local Similarity 28.0%; Pred. No. 2.5e-23;		
	Matches 81; Conservative 61; Mismatches 110; Indels 37; Gaps		
QY	60 NTILLPHIFTFVFSRKSIDMIISHLSHIHLLFTOAILVSLDFFGFSQNTODLRYK 119		
	!::! !: !::! !::! !: !::! !: !::! !: !::! !: !::! !: !::!		
Db	25 nsliffahcmfneenrskpdiclaflstqlmllvtng-liaadmfmaqgiwdlttr 83		
QY	120 VIVFLNKVMRGLSTCTPCLLSVLQAI-ISPISFLAKLKHPSASHILGFFLSWVL---- 174		

```

XX  WPI; 1997-245107/22.
DR  N-PSDB; AAT69548.
XX
PT  Nucleic acid molecule encoding vertebrate pheromone receptor -
PT  useful to identify modulators for control of reproductive and social
PT  behaviour, fertility and hormone secretion
XX
PS  Disclosure; Page 81-82; 123pp; English.
XX
CC  Rat pheromone receptors VN1, VN3, VN4, VN5, VN6, VN7 and VN2
CC  (AAW19103-09) are members of a novel family of presumed
CC  7-transmembrane domain receptors that are evolutionary independent
CC  of the odorant receptors of the main olfactory epithelium (MOE).
CC  Their amino acid sequences were deduced from cDNA clones (see also
CC  AAT69545-50, VN2 sequence not provided) obtd. by differential
CC  cloning, PCR and hybridisation from single vomeronasal sensory
CC  neurons (VSN). Individual neurons express different complements of
CC  the receptors. VN polypeptides can be expressed in host-vector
CC  systems for use in identifying modulators for control of maternal,
CC  reproductive and social behavior, to increase fertility, control
CC  hormone secretion and to regulate food uptake in humans and
CC  animals.
XX
SQ  Sequence 278 AA;

      Query Match      15.4%; Score 286.5; DB 18; Length 278;
      Best Local Similarity 28.1%; Pred. No. 3.6e-23;
      Matches 81; Conservative 57; Mismatches 109; Indels 41; Gaps 7;

QY  67 HFTFTVSHRSKIDMIISHLIHLLFTQAILVSLDFFGSQNTQDLRYKVIVLNLK 126
Db  3 hlfmifeknrskpidlyafislgtlmlitig-liaadmfmgrwdsttcqslilydr 61

QY  127 VNRLSICTPCLLSVLQAI-TPSPISFLAKLHPKPSASHLGFLEFSWLNMFIGN-IFCC 184
Db  :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
62 lirgftlcatclnlvltitispsscittfkshpsghlsafllfcvlyisfgshlfls 121

QY  185 TLRLP-----PVRGSSVCHTALFLFAHELHPQETVFTHTNDFEGCHL 227
Db  :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
122 tiatpnltsdnfmyvtksfcflpmsysrtsmfstpmal--realigliglssgymvaf 179

QY  228 YRVHGLKRLHGDYFIQIRGVLSTQACPRSPVVRASQAILLVSVFTYVWDTTF 287
Db  :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
180 wrhknqarhlh-----stlskskvspedratrtimlmsffvilylenv 224

QY  288 SPSGGVTWINDSLLVQLVIVANSAAISPLMLIYADNQIFKTLQMLW 335
Db  :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
225 vlysmftkdgamfcvqdlivshsyatispfvictetriik-----lw 268

RESULT 7
AAW19105
ID  AAW19105 standard; Protein; 310 AA.
XX
AC  AAW19105;
XX
DT  26-AUG-1997 (first entry)
XX
DE  Rat pheromone receptor VN4.
XX
KW  Pheromone receptor; vomeronasal sensory neuron; social behaviour;
KW  maternal behaviour; reproductive behaviour; fertility;
KW  hormone secretion.
XX
OS  Rattus sp.
XX
FH  Key Location/Qualifiers
FT  Domain 16..37
FT  /label= I
FT  /note= "predicted position of transmembrane domain
FT  I"

```

```

FT  Domain 50..71
FT  /label= II
FT  /note= "predicted position of transmembrane domain
FT  II"
FT  Domain 97..116
FT  /label= III
FT  /note= "predicted position of transmembrane domain
FT  III"
FT  Domain 135..156
FT  /label= IV
FT  /note= "predicted position of transmembrane domain
FT  IV"
FT  Domain 194..212
FT  /label= V
FT  /note= "predicted position of transmembrane domain
FT  V"
FT  Domain 240..258
FT  /label= VI
FT  /note= "predicted position of transmembrane domain
FT  VI"
FT  Domain 270..292
FT  /label= VII
FT  /note= "predicted position of transmembrane domain
FT  VII"
FT  Misc-difference 261
FT  /note= "Ser-261 is additional to the amino acid
FT  residues deduced from the nucleotide
FT  sequence"
FT  Misc-difference 263..264
FT  /note= "the amino acid sequence deduced from the
FT  nucleotide sequence has a lysine residue
FT  between Met-263 and Phe-264"
PN  WO9714790-A1.
XX
PD  24-APR-1997.
XX
PF  18-OCT-1996; 96WO-US16637.
XX
PR  19-OCT-1995; 95US-0005698.
XX
PA  (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI  Axel R, Dulac C;
XX
DR  WPI; 1997-245107/22.
DR  N-PSDB; AAT69547.
XX
PT  Nucleic acid molecule encoding vertebrate pheromone receptor -
PT  useful to identify modulators for control of reproductive and social
PT  behaviour, fertility and hormone secretion
XX
PS  Disclosure; Page 79-80; 123pp; English.
XX
CC  Rat pheromone receptors VN1, VN3, VN4, VN5, VN6, VN7 and VN2
CC  (AAW19103-09) are members of a novel family of presumed
CC  7-transmembrane domain receptors that are evolutionary independent
CC  of the odorant receptors of the main olfactory epithelium (MOE).
CC  Their amino acid sequences were deduced from cDNA clones (see also
CC  AAT69545-50, VN2 sequence not provided) obtd. by differential
CC  cloning, PCR and hybridisation from single vomeronasal sensory
CC  neurons (VSN). Individual neurons express different complements of
CC  the receptors. VN polypeptides can be expressed in host-vector
CC  systems for use in identifying modulators for control of maternal,
CC  reproductive and social behavior, to increase fertility, control
CC  hormone secretion and to regulate food uptake in humans and
CC  animals.
XX
SQ  Sequence 310 AA;

      Query Match      14.2%; Score 263.5; DB 18; Length 310;
      Best Local Similarity 26.4%; Pred. No. 1.5e-20;

```


FT Domain /label- Transmembrane_domain
FT 182..199
FT /label- Transmembrane_domain
FT 235..254
FT /label- Transmembrane_domain
XX
FN WO200157085-A2.
XX
XX 09-AUG-2001.
XX
XX 01-FEB-2001; 2001WO-US03455.
XX
XX 02-FEB-2000; 2000US-0180093.
XX 11-FEB-2000; 2000US-0182045.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Baughn MR, Au-Young J, Yue H;
XX
XX WPI; 2001-488869/53.
XX N-PSDB; AAD12956.
XX
XX Novel isolated human G-protein coupled receptor useful for diagnosing,
PT preventing and treating cell proliferative, neurological,
PT cardiovascular, gastrointestinal, autoimmune/inflammatory and metabolic
PT disorders -
XX
XX Claim 1; Page 121; 138pp; English.
XX
XX The present sequence is human G-protein coupled receptor-13 (GCRC-13)
CC protein. The present invention relates to GCRC protein and nucleic
CC acids encoding them. GCRC protein, its agonist or antagonist are useful
CC for treating diseases or conditions associated with decreased expression
CC or overexpression of functional GCRC in a patient, where the disorder
CC is selected from cell proliferative disorders such as actinic keratosis,
CC arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, psoriasis, and
CC cancer, neurological disorders such as epilepsy, stroke, Alzheimer's
CC disease, Huntington's disease, Parkinson's disease, cardiovascular
CC disorders such as hypertension, vasculitis, varicose veins, gastro-
CC intestinal disorders such as dysphagia, dyspepsia, anorexia, nausea,
CC pancreatitis, autoimmune/inflammatory disorders such as acquired
CC immunodeficiency syndrome (AIDS), Addison's disease, Crohn's disease,
CC uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic
CC infections, trauma and metabolic disorders such as diabetes, obesity,
CC osteoporosis. GCRC proteins and their cDNAs are used to assess the
CC effects of exogenous compounds on the expression of GCRC sequences.
CC GCRC cDNA is useful to create knock in humanised animals (pigs) or
CC transgenic animals (mice or rats) to model human disease, for
CC therapeutic or diagnostic purposes, for somatic or germline gene
CC therapy, to generate hybridisation probes useful in mapping the
CC naturally occurring genomic sequence, and in molecular biological
CC techniques.
XX
XX Sequence 254 AA;
XX
XX Query Match 10.7%; Score 198.5; DB 22; Length 254;
XX Best Local Similarity 25.7%; Pred. No. 1.8e-13;
XX Matches 70; Conservative 42; Mismatches 91; Indels 69; Gaps 7;
XX
XX 52 QASFGISANTILLPHITTFVFSRHSKIDMIISHLHILHLLFTQAILVSLDFGGSQN 111
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
XX 14 qtvvgvlgsvfvlhlysfctgcrldtdlvkhllvanflairckvqptmaafgvy 73
XX | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
XX 112 TODDRYKIVFVFNKVMRGLSCTPCLLSVLQAI-IPSFISLAKLKPSSASHILGFFLF 170
XX | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 74 finalgcklvfvlhrgvrgvsgigtcllsvfqvltvssrkrwklkpkhgvfsvll 133
XX | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
XX 171 SWLNMNFGVFCCLRLPPVKRGSCVCHTFALFLFAHELHPQETVFTHTNDFEGCHLYRV 230
XX | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 134 cwivcmvlvniif-----pmvyvtgkwn-----ytn-----itv 160
XX
XX 231 HCPKRLHG---DPIQIIRGVSFAFTQAC----- 258

Db 161 nedlgycsggnknkiagtllramllsfpdvclclgmlwvssmvcilhrhkqrghdrsd 220
QY 259 --PRVSPVKRASQAILLLVS-FVFTYVWDFTF 287
Db 221 lsraspenratqslilivstfvssytiscif 252
RESULT 10
AAB97828
ID AAB97828 standard; Protein; 394 AA.
XX
XX AAB97828;
XX
XX 09-AUG-2001 (first entry)
XX
XX Human G protein-coupled receptor PFI-014 protein sequence SEQ ID NO:2.
XX
XX Human; G protein-coupled receptor; GPCR; PFI-014; anorectic; cytostatic;
KW antidiabetic; neuroprotective; antiinflammatory; vulnery; osteopathic;
KW dermatologic; antimicrobial; antiallergic; obesity; diabetes; cancer;
KW metabolic disease; neurological disease; psychotherapeutic; inflammation;
KW urogenital disease; reproduction; sexual medicine; tissue repair;
KW dermatology; skin pigmentation; photoageing; frailty; osteoporosis;
KW cardiovascular disease; gastrointestinal disease; antinfection; allergy;
KW respiratory disease; sensory organ disorder; sleep disorder; hair loss.
OS Homo sapiens.
XX
XX EPI096008-A2.
XX
XX 02-MAY-2001.
XX
XX 16-OCT-2000; 2000EP-0309078.
XX
XX 22-OCT-1999; 99GB-00251115.
XX (PFI2) PFIZER LTD.
XX (PFI2) PFIZER INC.
XX
XX Hernandez R;
XX
XX WPI; 2001-309853/33.
XX N-PSDB; AAH20183.
XX
XX New human G-protein coupled receptor (GPCR) polypeptide, PFI-014 for
PT raising antibodies useful for treating obesity, neurological diseases,
PT diabetes, metabolic diseases, inflammation, cancer and dermatological
PT diseases -
XX
XX Claim 22; Page 36; 41pp; English.
XX
XX The present sequence represents a human G-protein coupled receptor (GPCR)
CC like protein, designated PFI-014. PFI-014 has anorectic, antidiabetic,
CC neuroprotective, antiinflammatory, cytostatic, vulnery, dermatological,
CC osteopathic, antimicrobial and antiallergic activities. Antibodies (Ab)
CC against PFI-014, compounds (C) which modulate PFI-014 and pharmaceutical
CC compositions (PC) comprising (Ab) or (C) are useful as pharmaceuticals.
CC (C) are useful in the manufacture of a medicament for the treatment of a
CC patient having the need to modulate i.e., agonise, antagonise or
CC selectively antagonise PFI-014. (Ab) and (C) are useful in the treatment
CC of a patient having a need to modulate PFI-014. (Ab), (PC) and (C) are
CC useful for treating obesity, diabetes and metabolic disease, neurological
CC disease, psychotherapeutics, urogenital disease, reproduction and sexual
CC medicine, cancer, inflammation, tissue repair, dermatology, skin
CC pigmentation, photoageing, frailty, osteoporosis, cardiovascular disease,
CC gastrointestinal disease, antinfection, allergy and respiratory disease,
CC sensory organ disorders, sleep disorders and hair loss.
XX
XX Sequence 394 AA;
XX
XX Query Match 10.4%; Score 193.5; DB 22; Length 394;

Best Local Similarity 25.3%; Pred. No. 1.2e-12;
Matches 75; Conservative 55; Mismatches 123; Indels 43; Gaps 10;

QY 52 QASFGISANTILLFHFTEVFSHRKSIDMIISHLSLHILLFTQAILVSLDFEGSON 111
Db 96 qvvvgllgnfsllyymfrygkprstdlilrhltvadsilvlskripetmatgikh 155

QY 112 TQDRLRYKVIVFLNKMVRLGSLCTPCLLSVLOAI-ISPSTFSLAKLKHPSASHILGFFLF 170
Db 156 fdnyfgckfilyahrvgrvsgistcellsfvqitnprsrwaemkvkaptvylgnil 215

QY 171 SWVLNMFIVGFCTLRLPPV-----KRGSSVCHTALF-LFAHELHPQETVF 217
Db 216 cwfahmlvnaif-----plyttgkwnnnitckgdgycsaplsdevcksvyaaltsf 268

QY 218 HTNDFEGCHLYRVHGPLKRL--HGDVFIQIRGVLSAFTQACPRVSPVKRASQAILLV 275
Db 269 hdlvclglmlwassivlvlyrhkqgvqhcrnll-----ypnspgnraigsllalv 321

QY 276 SFVFTYVWDTFFSGGVTTWIN-----DSLVLWL---QVIVANSYAAISPLMLIYAD 324
Db 322 s---tfaalcysalf---ityvyalfdnsswlvntaalliactpftisfpvlmcrd 371

RESULT 11
ID AAG64295 standard; Protein; 353 AA.
XX AAG64295;
XX 21-SEP-2001 (first entry)
XX Human GTP-binding protein-coupled receptor GPRV31.
XX GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
XX muscular; urinary; circulatory; anorectic; human; guanosine triphosphate;
XX G-protein.
XX Homo sapiens.
XX W0200148189-A1.
XX 05-JUL-2001.
XX 28-DEC-2000; 2000WO-JP09409.
XX 28-DEC-1999; 99JP-0375152.
XX 31-MAR-2000; 2000JP-0101339.
XX 23-MAY-2000; 2000JP-0155978.
XX (HELI-) HELIX RES INST.
XX Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;
XX Sugiyama T;
XX WPI; 2001-425663/45.
XX N-PSDB; AAH49508.
XX Family of guanosine triphosphate binding protein coupled receptors and
XX genes encoding them for treatment and prevention of diseases associated
XX with these receptors -
XX Claim 1; Pages 78-80; 137pp; Japanese.
XX The present sequence is the protein sequence for a human guanosine
XX triphosphate (GTP)-binding protein-coupled receptor. The receptor is
XX useful for the investigation, diagnosis, treatment and prevention of
XX diseases associated with GTP-binding protein-coupled receptors, including
XX neurological, circulatory, digestive system, immune system, muscle and
XX urinary system disorders. GTP-binding proteins are also known as
XX G-proteins.
XX Sequence 353 AA;

Query Match 10.4%; Score 192.5; DB 22; Length 353;
Best Local Similarity 24.8%; Pred. No. 1.3e-12;
Matches 91; Conservative 55; Mismatches 140; Indels 81; Gaps 13;

QY 2 LKLVIIENMAEIMLFSLDOLLFTDILCFN---FPSMKIKLP-GFITIQIFFYYPQASFGI 57
Db 6 lkl-----lslmtryffllfystdsdlnenqphldfdemafgkvksgisfligtvgv 60

QY 58 SANTILLLHFIFVFSHRKSIDMIISHLSLHILLFTQAILVSLDFEGSQNTQDDLR 117
Db 61 lgnsfllcfnlllftghklrptdlilsqalansmvlfkqipqtnaafgkylindtg 120

QY 118 YKIVFLNKMVRLGSLCTPCLLSVLOAI-ISPSTFSLAKLKHPSASHILGFFLFSSVLNM 176
Db 121 ckfvfyhrrvgrvsticllngfqaiknpsicrmeikirsprfi----- 168

QY 177 FIGVFCCTLRLPPVKGQSSVCHTALF-LFAHELHPQETVFHTNDFEGCHLYRVHGPLKR 236
Db 169 ----dfccllcwaphvlmnasv-----lllvngplnksnssaknn--ygycsykaskrfss 218

QY 237 LHG-----DY----FIQTIRGYLSAF-----TOPACPRVSPVKRASQAILL 273
Db 219 lhavlyfspdflmgfmvwasgsmvfflyrhkqgvqhnhnsrlsc-rpsgearathtimv 277

QY 274 LVSFVFTYVWDTFFSGGVTTWINDSLVLWL-----QVIVANSYAAISPLML 320
Db 278 lvssffvfyvshf-----ltiwtvtvngqgvivtnsvlvascfparspflv 325

QY 321 IYADNQI 327
Db 326 imsdthi 332

RESULT 12
AAB20349 standard; Protein; 353 AA.
XX AAB20349;
XX 11-JUN-2001 (first entry)
XX Human vomeronasal-like receptor hVLR1 (long form).
XX Vomeronasal-like receptor; hVLR1; V1RL1; human; pheromone;
XX G protein coupled receptor; signal transduction.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Domain 51..66
XX /note= "transmembrane domain 1"
XX Domain 85..107
XX /note= "transmembrane domain 2"
XX Domain 133..152
XX /note= "transmembrane domain 3"
XX Domain 172..189
XX /note= "transmembrane domain 4"
XX Domain 229..247
XX /note= "transmembrane domain 5"
XX Domain 274..293
XX /note= "transmembrane domain 6"
XX Domain 304..321
XX /note= "transmembrane domain 7"
XX W0200125431-A1.
XX 12-APR-2001.
XX 29-SEP-2000; 2000WO-US27211.
XX 01-OCT-1999; 99US-0157267.

```
PR 16-AUG-2000; 2000US-0225543.
PR 16-AUG-2000; 2000US-0640209.
PA (UYRQ ) UNIV ROCKEFELLER.
PI Mombaerts P, Rodriguez I;
XX N-PSDB; AAF30598.
XX
PT New primate, particularly human vomeronasal-like receptor, homologous
PT to rat and mouse pheromone receptors, useful to screen for
PT (ant)agonists and to identify receptor subtype selective ligands -
XX
PS Claim 2; Fig 1; 82pp; English.
XX
CC The present sequence is that of the long isoform of Caucasian human
CC vomeronasal-like receptor (hVLR1 or VLR1), a G protein coupled
CC receptor that is homologous to rat and mouse vomeronasal receptors
CC (putative pheromone receptors). A shorter isoform (see AAB20348),
CC expressed from an alternative downstream in-frame ATG start site,
CC and 2 variants (see AAB20350 and AAB20351) of hVLR1 were also
CC identified. hVLR1 functionality includes: binding pheromone-like
CC substrates and pheromone-like antagonists, which can be affected by
CC GTP; G-protein binding; and signal transduction in response to
CC binding vomeropherin or a vomeropherin agonist. hVLR1 mRNA
CC expression was found consistently in the olfactory mucosa. Methods
CC are provided for producing hVLR1, detecting expression of these
CC receptors, and screening assays for hVLR1 agonists and antagonists.
CC hVLR1 can also be used to identify critical ligand binding domains,
CC to determine selectivity of ligands, and to investigate signal
CC transduction systems of vomeronasal-like receptors.
XX
SQ Sequence 353 AA;

Query Match 10.4%; Score 192.5; DB 22; Length 353;
Best Local Similarity 24.8%; Pred. No. 1.3e-12;
Matches 91; Conservative 55; Mismatches 140; Indels 81; Gaps 13;

QY 2 LKLVIIENMAETLFLSDLLFSTDTLCFN---FPSKMIKLP-GFTITQIFFYPOASRGI 57
Db ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
6 lkl-----lslmtryffllfytdssdlnenqhlpdfdemafigkvksgisfligtvgi 60
QY 58 SANTILLPHITFTVFSHRKSIDMIISHLIHLITQAILVSLDFFGSGQNTQDDLRL 117
Db || : : : || : : : || : : : || : : : || : : : || : : : || : : : ||
61 lgnstllcfynllltghklrptdlilsqalansmvllfkqipqtaamafgikylindtg 120
QY 118 YKVIVFLNKMVRLSICTPCLLSVLQAI-ISPSIESLAKLKHPSASHILGFLFSWVLNM 176
Db || : : : || : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
121 cxfvfyhrvgrtvssticlingfgaiklnpsicrweikirsprfl----- 168
QY 177 FIGVIFCCPLRPPVKRGOSVCHTALFAHELHPQETVFTHTNDFEGCHLYRVHGPLKR 236
Db ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
169 -----dfcclllcwaphvlmasv-----lllvngplnskassaknn--yyycsykaskrfss 218
QY 237 LHG-----DY----FIQIRYLSAF-----TQPACPRVSPVKRASQAILL 273
Db || : : : || : : : || : : : || : : : || : : : || : : : || : : : ||
219 lhaviyspdmfslgmfwagsmvfflyrhkqvgqhnhsnrlsc-rpsqearathtimv 277
QY 274 LVSFVFTYVWDFTFSESGVTWINDSLVLWL-----QVIIVANSYAALISPLML 320
Db || : : : || : : : || : : : || : : : || : : : || : : : || : : : ||
278 lvssffvysvhsf-----ltiwtvtvanpgqwiavtnsvlvascparspfv 325
QY 321 IYADNQI 327
Db : : : || : : : || : : : || : : : || : : : || : : : || : : : ||
326 lmsdthi 332

RESULT 13
AAB20350
ID AAB20350 standard; Protein; 313 AA.
XX
```

```
AC AAB20350;
XX
DT 11-JUN-2001 (first entry)
XX
DE Human vomeronasal-like receptor variant VLR1la (S201F/A229D).
XX
KW Vomeronasal-like receptor; hVLR1; VLR1la; human; pheromone;
KW G protein coupled receptor; signal transduction; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 201 /note= "site of S201F substitution"
FT FT Misc-difference 229 /note= "site of A229D substitution"
FT FT Domain 11..26 /note= "transmembrane domain 1"
FT FT Domain 45..67 /note= "transmembrane domain 2"
FT FT Domain 93..112 /note= "transmembrane domain 3"
FT FT Domain 132..149 /note= "transmembrane domain 4"
FT FT Domain 189..207 /note= "transmembrane domain 5"
FT FT Domain 234..253 /note= "transmembrane domain 6"
FT FT Domain 264..281 /note= "transmembrane domain 7"
XX
PN WO200125431-A1.
XX
PD 12-APR-2001.
XX
XX 29-SEP-2000; 2000WO-US27211.
XX
PR 01-OCT-1999; 99US-0157267.
PR 16-AUG-2000; 2000US-0225543.
PR 16-AUG-2000; 2000US-0640209.
XX
XX (UYRQ ) UNIV ROCKEFELLER.
XX
PI Mombaerts P, Rodriguez I;
XX
DR WPI; 2001-273578/28.
XX
PT New primate, particularly human vomeronasal-like receptor, homologous
PT to rat and mouse pheromone receptors, useful to screen for
PT (ant)agonists and to identify receptor subtype selective ligands -
XX
PS Claim 3; Fig 7a-b; 82pp; English.
XX
CC The present sequence is that of variant VLR1la of human
CC vomeronasal-like receptor (hVLR1, see AAB20348). The variant has
CC Ser substituted by Phe at amino acid position 201, and Ala
CC substituted by Asp at amino acid position 229. The amino acid
CC substitutions are the result of 2 single nucleotide polymorphisms.
CC S201F, A229D and S201F/A229D allelic variants have been identified.
CC VLR1la is found in Indonesian, Pygmy, Japanese, Cambodian,
CC Amerindian and Adaygei subjects, and hVLR1 in Caucasians. hVLR1
CC (VLR1l) shows homology to rat and mouse pheromone receptors. Its
CC functionality includes: binding pheromone-like substrates and
CC pheromone-like antagonists, which can be affected by GTP; G-protein
CC binding; and signal transduction in response to binding vomeropherin
CC or a vomeropherin agonist. Methods are provided for producing
CC hVLR1, detecting expression of these receptors, and screening
CC assays for hVLR1 agonists and antagonists. hVLR1 can also be used
CC to identify critical ligand binding domains, to determine
CC selectivity of ligands, and to investigate signal transduction
CC systems of vomeronasal-like receptors.
XX
SQ Sequence 313 AA;
```

```

Query Match      10.4%; Score 192; DB 22; Length 313;
Best Local Similarity 24.7%; Pred. No. 1.2e-12;
Matches 80; Conservative 50; Mismatches 122; Indels 72; Gaps 10;

QY 41 GFTTIQIFYPQASFGISANTILLFHFVFSHRKSIDMTIISHLSLHILLFTQAI 100
DB 4 gkvksqisfliqtvgilgnslfllcfnlllftghklrptdlilsglalansmvlffkgl 63
QY 101 LVSLDFEGSQNTODDLRYKVVIVLNKVMRGLSICTPCLLSLVQAI-ISPSIFSIAKLKHP 159
DB 64 pqumaafgikyllndtgcxfvyhrvtrvsticllingfqaiklncpsicrwmelkir 123
QY 160 SASHILGFFLFSWVLMNFGVFCCTLRPLPPVVRKGSSVCHTALFLFAHELHPQETVFHT 219
DB 124 sprfi-----dfccllcwaphvlmasv-----lllvngplnksnssakn 163
QY 220 NDFEGCHLYRVHGPKRLKHG-----DY-----FIQIRGYLSAF-----TQP 256
DB 164 n--ygycsykaskrfslhavlyfspdmslghfmvwasgfmvfflyrhkqgvqnhnsnrl 221
QY 257 ACPRVSPVKRASQAIIILLVSFVTYVDETFPSFGGVTWINDSLLYWL----- 304
DB 222 sc-rpsqdrathlmlvssffvfysvhsf-----ltiwtvvanpgqvwt 268
QY 305 -QVIVANSYAAISPLMIYADNOI 327
DB 269 nsvlvasciparspflmsdthl 292

RESULT 14
AAB20351
XX AAB20351 standard; Protein; 313 AA.
XX
AC AAB20351;
XX
XX 11-JUN-2001 (first entry)
XX
DE Human vomeronasal-like receptor variant (S201F).
XX
XX Vomeronasal-like receptor; hVLR1; VLR1; human; pheromone;
KW G protein coupled receptor; signal transduction; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Misc-difference 201
FT /note= "site of S201F substitution"
FT Domain 11..26
FT /note= "transmembrane domain 1"
FT Domain 45..67
FT /note= "transmembrane domain 2"
FT Domain 93..112
FT /note= "transmembrane domain 3"
FT Domain 132..149
FT /note= "transmembrane domain 4"
FT Domain 189..207
FT /note= "transmembrane domain 5"
FT Domain 234..253
FT /note= "transmembrane domain 6"
FT Domain 264..281
FT /note= "transmembrane domain 7"
XX
XX WO200125431-A1.
XX
XX 12-APR-2001.
XX
XX 29-SEP-2000; 2000WO-US27211.
XX
XX 01-OCT-1999; 99US-0157267.
PR 16-AUG-2000; 2000US-0225543.
PR 16-AUG-2000; 2000US-0640209.
```

```

XX (UYRQ ) UNIV ROCKEFELLER.
XX
XX Mombaerts P, Rodriguez I;
XX
XX WPI; 2001-273578/28.
XX
XX New primate, particularly human vomeronasal-like receptor, homologous
PT to rat and mouse pheromone receptors, useful to screen for
PT (antagonists and to identify receptor subtype selective ligands -
XX
XX Claim 3; Page -: 82pp; English.
XX
XX The present sequence is that of a variant of human vomeronasal-like
CC receptor (hVLR1 or VLR1, see AAB20348). The variant has Ser
CC substituted by phe at amino acid position 201. S201F, A229D and
CC S201F/A229D allelic variants have been identified. The amino acid
CC substitutions are the result of single nucleotide polymorphisms.
CC The S201F substitution was detected in Amerindian and Adygei
CC subjects. hVLR1 shows homology to rat and mouse pheromone receptors.
CC Its functionality includes: binding pheromone-like substrates and
CC pheromone-like antagonists, which can be affected by GTP; G-protein
CC binding; and signal transduction in response to binding vomeropherin
CC or a vomeropherin agonist. Methods are provided for producing
CC hVLR1, detecting expression of these receptors, and screening
CC assays for hVLR1 agonists and antagonists. hVLR1 can also be used
CC to identify critical ligand binding domains. hVLR1 can also be used
CC selectivity of ligands, and to investigate signal transduction
CC systems of vomeronasal-like receptors.
CC Note: the present sequence is not shown in the specification but is
CC derived from the hVLR1 sequence given in Figure 1 (see AAB20348).
XX
XX Sequence 313 AA;

Query Match      10.4%; Score 192; DB 22; Length 313;
Best Local Similarity 24.7%; Pred. No. 1.2e-12;
Matches 80; Conservative 50; Mismatches 122; Indels 72; Gaps 10;

QY 41 GFTTIQIFYPQASFGISANTILLFHFVFSHRKSIDMTIISHLSLHILLFTQAI 100
DB 4 gkvksqisfliqtvgilgnslfllcfnlllftghklrptdlilsglalansmvlffkgl 63
QY 101 LVSLDFEGSQNTODDLRYKVVIVLNKVMRGLSICTPCLLSLVQAI-ISPSIFSIAKLKHP 159
DB 64 pqumaafgikyllndtgcxfvyhrvtrvsticllingfqaiklncpsicrwmelkir 123
QY 160 SASHILGFFLFSWVLMNFGVFCCTLRPLPPVVRKGSSVCHTALFLFAHELHPQETVFHT 219
DB 124 sprfi-----dfccllcwaphvlmasv-----lllvngplnksnssakn 163
QY 220 NDFEGCHLYRVHGPKRLKHG-----DY-----FIQIRGYLSAF-----TQP 256
DB 164 n--ygycsykaskrfslhavlyfspdmslghfmvwasgfmvfflyrhkqgvqnhnsnrl 221
QY 257 ACPRVSPVKRASQAIIILLVSFVTYVDETFPSFGGVTWINDSLLYWL----- 304
DB 222 sc-rpsqdrathlmlvssffvfysvhsf-----ltiwtvvanpgqvwt 268
QY 305 -QVIVANSYAAISPLMIYADNOI 327
DB 269 nsvlvasciparspflmsdthl 292

RESULT 15
AAB20348
ID AAB20348 standard; Protein; 313 AA.
XX
XX AAB20348;
XX
XX 11-JUN-2001 (first entry)
XX
XX Human vomeronasal-like receptor hVLR1 (short form).
```

XX Vomeronasal-like receptor; hVLR1; VLR1; human; pheromone;
 KW G protein coupled receptor; signal transduction.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH Domain 11..26
 FT Domain /note= "transmembrane domain 1"
 FT Domain 45..67
 FT Domain /note= "transmembrane domain 2"
 FT Domain 93..112
 FT Domain /note= "transmembrane domain 3"
 FT Domain 132..149
 FT Domain /note= "transmembrane domain 4"
 FT Domain 189..207
 FT Domain /note= "transmembrane domain 5"
 FT Domain 234..253
 FT Domain /note= "transmembrane domain 6"
 FT Domain 264..281
 FT Domain /note= "transmembrane domain 7"

XX WO200125431-A1.

XX PD 12-APR-2001.

XX 29-SEP-2000; 2000WO-US27211.

XX 01-OCT-1999; 99US-0157267.

XX 16-AUG-2000; 2000US-0225543.

XX 16-AUG-2000; 2000US-0640209.

XX (UVRQ) UNIV ROCKEFELLER.

XX Mombaerts P, Rodriguez I;

XX WPI; 2001-273578/28.

XX N-PSDB; AAF30597.

XX New primate, particularly human vomeronasal-like receptor, homologous
 PT to rat and mouse pheromone receptors, useful to screen for
 PT (ant)agonists and to identify receptor subtype selective ligands -

XX Claim 2; Fig 1; 82pp; English.

XX The present sequence is that of the short isoform of Caucasian human
 CC vomeronasal-like receptor (hVLR1 or VLR1), a G protein coupled
 CC receptor that is homologous to rat and mouse vomeronasal receptors
 CC (putative pheromone receptors). A longer isoform (see AAB20349),
 CC expressed from an alternative upstream in-frame ATG start site,
 CC and 2 variants (see AAB20350 and AAB20351) of hVLR1 were also
 CC identified. hVLR1 functionality includes: binding pheromone-like
 CC substrates and pheromone-like antagonists, which can be affected by
 CC GTP; G-protein binding; and signal transduction in response to
 CC binding vomeropherin or a vomeropherin agonist. hVLR1 mRNA
 CC expression was found consistently in the olfactory mucosa. Methods
 CC are provided for producing hVLR1, detecting expression of these
 CC receptors, and screening assays for hVLR1 agonists and antagonists.
 CC hVLR1 can also be used to identify critical ligand binding domains,
 CC to determine selectivity of ligands, and to investigate signal
 CC transduction systems of vomeronasal-like receptors.

XX Sequence 313 AA;

Query Match 10.1%; Score 187; DB 22; Length 313;

Best Local Similarity 24.7%; Pred. No. 4.5e-12;

Matches 80; Conservative 49; Mismatches 123; Indels 72; Gaps 10;

QY 41 GPITTIQIFYPQASFGISANTILLFLHFTVFVSHRSKSIDMIHSLIHLILFTQAI 100

Db 4 gkvksqisflqtgvgilgnsfilcfynliiftghklrptdliisqlansmvlfkqi 63

QY 101 LVSLDFGSGNTQDDRLRYKVIIVFLNKNVRGLSICTPCLLSVLQAI-ISPSTFSIAKLKHP 159
 Db 64 pgtmaafglkylldtgckfvyhrvgtrvslsticllngfalklnpsicrwmeikir 123
 QY 160 SASHILGFFFLFSWVLNMFIVGFCCTLRLPVPVKRQSSVCHTALFLFAHELHPQETVFT 219
 Db 124 sprfi-----dfccllcwaphvlnnasv-----lllvngpnsksnsaakn 163
 QY 220 NDFEGCHLYRVHGPLKRLHG-----DY----FIQTIRGYLSAF-----TQP 256
 Db 164 n--ygyesykaskrfsslhavlyfspdfmslgfwnwasgsmvfflyrhkqgqhnhsnrl 221
 QY 257 ACPRVSPVKRASQAILLVSVFTYVWDFTFSFGGVTWINDSLVLW----- 304
 Db 222 sc-rpsgearathtimvlsvsfvfysvhsf-----ltiwtvtvvanpgqwiwt 268
 QY 305 -QVIVANSYAAISPLMLIYADNQI 327
 Db 269 nsvivascfparspfvlmsdthl 292

Search completed: September 18, 2002, 10:52:12
 Job time: 234 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 18, 2002, 10:48:38 ; Search time 13.07 Seconds
(without alignments)
667.172 Million cell updates/sec

Title: US-09-728-309-3

Perfect score: 1855

Sequence: 1 MLKLVIIENMAEIMFLSIDL.....YLSPPKMLKFNRCQGSTKK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	5.9	376	1	US-08-279-590A-2
2	109	5.9	376	2	US-08-910-092-2
3	103.5	5.6	503	1	US-08-484-840-4
4	103.5	5.6	503	1	US-08-483-094-4
5	93	5.0	384	1	US-07-937-609-17
6	93	5.0	384	4	US-08-029-170-17
7	93	5.0	384	5	PCT-US92-02091-2
8	92.5	5.0	502	1	US-08-484-840-3
9	92.5	5.0	502	1	US-08-483-094-3
10	89.5	4.8	384	5	PCT-US92-02091-3
11	88.5	4.8	674	4	US-07-757-342D-10
12	88.5	4.8	699	4	US-07-757-342D-2
13	85	4.6	245	4	US-08-845-258-32
14	85	4.6	245	4	US-08-990-571-32
15	85	4.6	245	4	US-08-723-142A-32
16	83	4.5	690	2	US-08-619-554-8
17	83	4.5	1394	4	US-09-213-053-2
18	82.5	4.4	297	1	US-07-866-560-4
19	82.5	4.4	297	1	US-08-077-673-4
20	82.5	4.4	297	1	US-08-671-525B-4
21	82.5	4.4	297	1	US-08-672-109B-4
22	82.5	4.4	297	1	US-08-478-992-4
23	82.5	4.4	297	2	US-08-842-045-4
24	82.5	4.4	297	2	US-08-842-238-4
25	82.5	4.4	297	2	US-08-780-749A-3
26	82.5	4.4	297	3	US-09-105-298-4
27	82.5	4.4	297	3	US-08-629-335B-4

28	82.5	4.4	297	4	US-08-870-511-3	Sequence 3, Appl1
29	82.5	4.4	328	1	US-08-118-270-42	Sequence 42, Appl1
30	82.5	4.4	328	5	PCT-US93-08528-42	Sequence 42, Appl1
31	82	4.4	823	1	US-07-745-206A-15	Sequence 15, Appl1
32	82	4.4	823	2	US-08-311-363-15	Sequence 15, Appl1
33	82	4.4	968	3	US-08-551-999A-7	Sequence 7, Appl1
34	82	4.4	968	4	US-09-385-752-7	Sequence 7, Appl1
35	82	4.4	1754	1	US-07-745-206A-13	Sequence 13, Appl1
36	82	4.4	1754	2	US-08-311-363-13	Sequence 13, Appl1
37	82	4.4	1876	2	US-08-619-554-2	Sequence 2, Appl1
38	82	4.4	2237	1	US-08-455-543A-48	Sequence 48, Appl1
39	82	4.4	2237	2	US-08-223-305C-48	Sequence 48, Appl1
40	82	4.4	2337	4	US-09-268-163-8	Sequence 8, Appl1
41	82	4.4	2337	3	US-08-713-118-2	Sequence 2, Appl1
42	82	4.4	2337	4	US-09-452-007-2	Sequence 2, Appl1
43	82	4.4	2339	1	US-08-455-543A-47	Sequence 47, Appl1
44	82	4.4	2339	2	US-08-223-305C-47	Sequence 47, Appl1
45	82	4.4	2339	4	US-09-268-163-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-08-279-590A-2
; Sequence 2, Application US/08279590A
; Patent No. 5656749
; GENERAL INFORMATION:
; APPLICANT: ELIOT R. SPINDEL, SRINIVASA
; APPLICANT: NAGALLA AND BRENDA BARRY
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE
; TITLE OF INVENTION: PEPTIDES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/279,590A
; FILING DATE: July 22, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: No. 5656749e
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Y. Rocky Tsao
; REGISTRATION NUMBER: 34,053
; REFERENCE/DOCKET NUMBER: 00537/098001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200134
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: N/A
; US-08-279-590A-2

Query Match 5.9%; Score 109; DB 1; Length 376;
Best Local Similarity 21.1%; Pred. No. 0.0021;
Matches 68; Conservative 49; Mismatches 120; Indels 86; Gaps 15;


```
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/166 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: mouse gastrin-releasing peptide receptor
US-07-937-609-17

Query Match          5.0%; Score 93; DB 1; Length 384;
Best Local Similarity 21.2%; Pred. No. 0.1;
Matches 73; Conservative 51; Mismatches 144; Indels 76; Gaps 13;

QY 40 PGFI-TIOIFFYQASFGISANTILLFHIFTVFESHRSKSIDMIISHLSLHILLFTQ 98
DB 38 PGFIVIPAVYGLIIVIGLIGN--ITLIKIFTVKSMRNP-NLFISLALGDLILLVTC 94
QY 99 AILVSLDFGSGONTODDLRYKIVIVFLNKMVRLGSLCTPCLLSV--LQAIISPSIFSLAKL 156
DB 95 APVDASKYLADRWLFGRIKGLIPFQILTSVGVSVFTLTALSADRYKAIVRPMDI----- 149
QY 157 KHPASHLIGFLFS-----WVLNMFIVGFCCTLRPP--VKRGSSVCHTALFLFAHEL 210
DB 150 ---QASHALMKICLKAALINIVSMLLAIPAEVFDLHPFHVKDTNQTFTISCAPYPHSNEL 206
QY 211 HPQETVFTHTNDFEGCHLYRVHG-----PLKRLHGDYFIQIRGYLSAFTQ 256
DB 207 HP-----KIHSMASFLVYVIPLAIIISVYFIARNLIQSAYNLPVEGN 250
QY 257 --ACPRVSPVKRASQAIIILLYSFVETYWVDTFSP---SGGVTVINDSLLVWLQIVIVANS 311
DB 251 IHVKQIESRKRLAKTVLVFVGLFAFCWLPNHHVILYRSYHYSEVDTSMLEHFTVSICAH 310
QY 312 YA----AISPLMLIYADNQIFKTLQMLWFKYLSPPKMLKENRQ 351
DB 311 LAFTNSCVNPFAL-----YLLSKSFRKQFNTQ 337

RESULT 6
US-08-029-170-17
; Sequence 17, Application US/08029170
; Patent No. 6169173
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; TITLE OF INVENTION: CLONECYSTOKININ RECEPTOR-ENCODING DNA
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/029,170
; FILING DATE: 19930310
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 07/937,609
; FILING DATE: 02-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/928,033
; FILING DATE: 11-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/861,769
; FILING DATE: 01-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,248
; FILING DATE: 07-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/166 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: mouse gastrin-releasing peptide receptor
US-08-029-170-17

Query Match          5.0%; Score 93; DB 4; Length 384;
Best Local Similarity 21.2%; Pred. No. 0.1;
Matches 73; Conservative 51; Mismatches 144; Indels 76; Gaps 13;

QY 40 PGFI-TIOIFFYQASFGISANTILLFHIFTVFESHRSKSIDMIISHLSLHILLFTQ 98
DB 38 PGFIVIPAVYGLIIVIGLIGN--ITLIKIFTVKSMRNP-NLFISLALGDLILLVTC 94
QY 99 AILVSLDFGSGONTODDLRYKIVIVFLNKMVRLGSLCTPCLLSV--LQAIISPSIFSLAKL 156
DB 95 APVDASKYLADRWLFGRIKGLIPFQILTSVGVSVFTLTALSADRYKAIVRPMDI----- 149
QY 157 KHPASHLIGFLFS-----WVLNMFIVGFCCTLRPP--VKRGSSVCHTALFLFAHEL 210
DB 150 ---QASHALMKICLKAALINIVSMLLAIPAEVFDLHPFHVKDTNQTFTISCAPYPHSNEL 206
QY 211 HPQETVFTHTNDFEGCHLYRVHG-----PLKRLHGDYFIQIRGYLSAFTQ 256
DB 207 HP-----KIHSMASFLVYVIPLAIIISVYFIARNLIQSAYNLPVEGN 250
QY 257 --ACPRVSPVKRASQAIIILLYSFVETYWVDTFSP---SGGVTVINDSLLVWLQIVIVANS 311
DB 251 IHVKQIESRKRLAKTVLVFVGLFAFCWLPNHHVILYRSYHYSEVDTSMLEHFTVSICAH 310
QY 312 YA----AISPLMLIYADNQIFKTLQMLWFKYLSPPKMLKENRQ 351
DB 311 LAFTNSCVNPFAL-----YLLSKSFRKQFNTQ 337

RESULT 7
PCT-US92-02091-2
; Sequence 2, Application PC/TUS9202091
; GENERAL INFORMATION:
; APPLICANT: Batteny Jr., James F.
; APPLICANT: Corjay, Martha H.
; APPLICANT: Feldman, Richard I.
; APPLICANT: Harkins, Richard N.
; TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edwin P. Ching
; STREET: 1501 Harbor Bay Parkway
; CITY: Alameda
; STATE: CA
```


APPLICANT: MOSCOW, Jeffrey A.
APPLICANT: COWAN, Kenneth H.
APPLICANT: DIXON, Kathy
APPLICANT: HE, Rui
TITLE OF INVENTION: A GENE ENCODING A HUMAN REDUCED FOLATE
CARRIER (RFC) AND METHODS FOR THE TREATMENT OF
TITLE OF INVENTION: METHOTREXATE-RESISTANT, TRANSPORT-DEFICIENT CANCER CELLS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,094
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/323/NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 502 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-483-094-3

Query Match 5.0%; Score 92.5; DB 1; Length 502;
Best Local Similarity 21.2%; Pred. No. 0.17;
Matches 82; Conservative 46; Mismatches 98; Indels 161; Gaps 19;
QY 49 FYPOASFGISANTILLPHITTFVSHRSKSIDMIISHLSLHILLFTQAILVSLDFPG 108
Db FMAQIRPGSEFTTPFLERKFT-----KQVNTNEIIPMLPYSHLAVLPVFLT----- 85
QY 109 SONTODDLRYKIVIFLN-----KVMRGLSICTPCLLSVLQAI-----ISPSIFS 152
Db DYLRYKPKVLLQCLSFVCVWMLLLILGTSVYHMQLMVEFVSVTMAARIATSYIFS 140
QY 153 LAKLKHP-----ASHI-----LGFFLFSVWL 174
Db LV-----HPSRYQRMASYRAANVLLGVFISVLGQALVTGCHISTYTLNCVSLGFLIFSLVL 197
QY 175 NMFIGVIFCCTLRLLPPVRR-----GOSSVCHTALFLFAHELHPQETVFHTNDFEGCHLYRV 230
Db SLF-----LKR-----KRSLFENSTLARGALPCELQDHP----- 229
QY 231 HGPLRLKHGDIYFIQIRG-----YLSATQPCRPVSPVKRASQAILLVSVFV----- 279
Db GPDRKL--DRMLGRCRDSFLVRMLSELVENA-----RQPQLRLCLMWWFNSSGY 278
QY 280 --TYW-----DTEFSFGGVV-----WINDSLLVNLOVIVANSYAAI 315
Db LITYYVHVLRWSTDSSLVNGAVDAASTLLSAITSFSGAGLSIRTWLSKLVIAQ 338
QY 316 SPLMLIYADNQIFKTLQW--LWFKYLS 340
Db ASL-----VFCMFQIRDIWVCVVT 357

RESULT 10
PCT-US92-02091-4
Sequence 4, Application PC/TUS9202091
GENERAL INFORMATION:
APPLICANT: Battley Jr., James F.
APPLICANT: Corjay, Martha H.
APPLICANT: Feldman, Richard I.
APPLICANT: Harkins, Richard N.
TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edwin P. Ching
STREET: 1501 Harbor Bay Parkway
CITY: Alameda
STATE: CA
COUNTRY: USA
ZIP: 94501
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02091
FILING DATE: 19920313
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/426,150
FILING DATE: 24-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/533,659
FILING DATE: 05-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34090
REFERENCE/DOCKET NUMBER: A-0092C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-266-7476
TELEFAX: 415-266-7400
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-02091-4

Query Match 4.8%; Score 89.5; DB 5; Length 384;
Best Local Similarity 21.3%; Pred. No. 0.24;
Matches 55; Conservative 37; Mismatches 103; Indels 63; Gaps 10;
QY 56 GISANTILLPHITTFVSHRSKSIDMIISHLSLHILLFTQAILVSLDFGSGNTQDD 115
Db GLIGH--ITLIKIFCTVKSMRNP-NLFISLALGDLLLLITCAPVDASRYLADRMLFGR 110
QY 116 LRYKIVIFLNKVMRGLSICTPCLLSV--LQALISPSIFSLAKLHPSASHILGFF----L 169
Db IGCKLPIQLTSGVSVFTLTALSADRYKAIVRPDI-----QASHALMKICLAA 162
QY 170 FSWLNMFIGVIFCCTLRLLPPVKRGQSSVCHTALFLFAHELHPQETVFHTNDFEGCHLY- 228
Db FIWISMLLAI-----PEAVFSDLHPFHEEST---NOTFISCAPYP 200
QY 229 -----RVHG-----PLKRLHGDYFIQTIRGLVSLAFTOP-----ACPRVSPVK 265
Db HSNLHPKIHSMASFLVFIYVPLSLIISVYFIYFIKLNLIQSAYNLPVEGNIHVKKOIESRK 260
QY 266 RASQAILLVSVFETYWY 283
Db RLAKTVLVFVGLFAFCWL 278

```

RESULT 11
US-07-757-342D-10
; Sequence 10, Application US/07757342D
; Patent No. 6218509
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Masao
; MINEGISHI, Takashi
; NAKAMURA, Kazuto
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,342D
; FILING DATE: 10-Sep-1991
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BUCKLEY, Linda M.
; REGISTRATION NUMBER: 31003
; REFERENCE/DOCKET NUMBER: 41226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 674 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-07-757-342D-10

Query Match 4.8%; Score 88.5; DB 4; Length 674;
Best Local Similarity 21.8%; Pred. No. 0.67;
Matches 84; Conservative 49; Mismatches 137; Indels 115; Gaps

QY 1 MLKLV1--TENMAEIMFLSLLDILLSTDLICNF-----PSKWKILPGGITIOIFYP 51
Db 50 VIKIEISQDLSERIDANFDNLLNSELIONTKNLYIEPGAFINLPGLKYLSI----- 105
QY 52 QASFGISANTILLRHIFTVFVSHRSKSIDMTIISHLSLIHLLFTQAI-----LVSLD 105
Db 106 -----CNTGIRKFPDVKVFESSESNFILEICDNL---HITIPGNAFQGMNESVTLK 155
QY 106 FFG-----SNTQODDLRYKVIVFLNKVMRG-----LSICPCL----- 138
Db 156 LYNGEFEQVSHAFNQTTLTSLKLNHVLEKMHNGAFRGATGPKTLDISSTKLQALPSY 215
QY 139 -LSVLQAITSPTSIFSLAKLHPKSASHILGFFLFSVNLNFTGVI-----FCTLRLP 189
Db 216 GLIESQRLATSYSLUKL--PSR-----ETFVNLLSATITYPSHCCAFNRL 260
QY 190 PVRRGQSSVCHTALFLFAHELHPQETVFTHTNDFEGCHLYRVHGPLKRLHGDFYIQT-IRG 248
Db 261 PTK--EQNFSHSISENFSKQC--ESTVRKVSN-----KTLYSNLAESLSG 303
QY 249 --YLSAFTQAPACRPVSPVKRASQAILLLVSFVF-----TYWVDFTFSFGGVTVINDSLLV 303
Db 304 WDYEYGFCLPKYPRCAPEADFNPCEDIMGYDFLRVLTLWLINILAIMGNMT-----V 355

```

```

Qy   304 LOIVANSYAAISPLMLI-----YAD 324
      | |: : | | | | | | | |
Db   356 LFLVLLTRYKLTVPRLMCLNSPAD 380

RESULT 12
US-07-757-342D-2
; Sequence 2, Application US/07757342D
; Patent No. 6218509
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Masao
;             NINEGISHI, Takashi
;             NAKAMURA, Kazuto
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;              CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,342D
; FILING DATE: 10-Sep-1991
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: EUCKLEY, Linda M.
; REGISTRATION NUMBER: 31003
; REFERENCE/DOCKET NUMBER: 41226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-07-757-342D-2

Query Match          4.8%; Score 88.5; DB 4; Length 699;
Best Local Similarity 21.8%; Pred. No. 0.71;
Matches 84; Conservative 49; Mismatches 137; Indels 115; Gaps

Qy       1 MLKLVII--IENMAEIMLFSLDLLSTFDIUCFNF-----PSKMIKLPGFITIQIFFYP 51
      :|:|:| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db       75 VKIEIQSDLSERIANAFDNLLNISEILIQNTKNRIEPCAFNLGLKYLSI---- 130

Qy       52 QASEGISANTILLFHIFTVFVSHRSKIDMIISHLSLIHLLFTQAII-----LYSLD 105
      || | |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db       131 -----CNTGIKFPDVTKVFSSSENFLEICDNL---HITTPGNATFGMNNESTVLK 180

Qy       106 FFG-----SQNTQDDLRYKVIVFLNKVMRG-----LSICTPCL----- 138
      :| | |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db       181 LYNGNGFEVOASHAFNGTTLTSLSELKENVHLEKMHNGAFRGATGPRTLDISTSKLALPSY 240

Qy       139 -LSVLQAIIGSPISFLAKLKHPSASHILGTFELSWLNMFIVI-----FCCTLRUP 189
      || | |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db       241 GLESIQLRIATSSYSKLK--PSR-----ETPVNLEATLYPSHCACFRML 285

Qy       190 PVKRGSQSVCHTALFLFAHELHPQETVFHTNDPECGCHLYRVHGFLKRLHGDYFIQT--IRG 248
      || | |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

```

```

Query Match      4.8%; Score 88.5; DB 4; Length 699;
Best Local Similarity 21.8%; Pred. No. 0.71;
Matches 84; Conservative 49; Mismatches 137; Indels 115; Gaps 19;

Qy      1  MLKVI--IENNAIMLFSLDLLFSDIILCFNF-----PSKNIKLPQFIITQIFYP 51
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      75  VKIEIQSDLSERIEANAFDNLNLSEITQNTKRLRIEPGAFINLGLKYLSI---- 130
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy     52  QASGFSANTILLFHFIITVFVSHRSKIDMIISHLSLHILLFTQAI-----LYSLD 105
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     131  -----CNTGIRKFPDVKVFSSSNFLEICDNL---HITTPGNAFGMNNESVTLK 180
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy     106  FFG-----SQNTQDRLRVKIVFLNKVMRG-----LSICTPCL----- 138
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     181  LYGNFVEVQSHAFNGTTLTSLLEKENVHLEKMHNGAFRGATGPKTLDISTKLAQLPSY 240
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy     139  -LSVLQAIISPIESLAKLKHPSASHILGFFLFSWLNMFIGVI-----FCTLRLP 189
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     241  GLESIQRLIATSSYSLKLL--PSR-----ETFNVLLEATLTPSHCCAERNL 285
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy     190  PVKRQSSVCHTALFLFAHELHPQETVHTNDFEGCHLYRVGHGPKRLGHGDIYQIOT--IRG 248
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 286 PTK--EQNFHSISNFSKQC--ESTVRKVN-----KTLYSMLAESELSG 328
QY 249 --YLSAFTQACPRVSPVKRASQAILLVSVF---TYWVDFTFSGVGTWINDSLVW 303
Db 329 WDYEGFCGLPKTPRCAPEDAFNPCEIDMGDFLRVLWLINILAINGNMT-----V 380
QY 304 LQVIVANSYAAISPLMLI-----YAD 324
Db 381 LFLVLTSLRYKLTVPFELMCNLSFAD 405

RESULT 13
US-08-845-258-32
; Sequence 32, Application US/08845258
; Patent No. 6183976
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,258
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.426C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-845-258-32

Query Match 4.6%; Score 85; DB 4; Length 245;
Best Local Similarity 22.4%; Pred. No. 0.39;
Matches 68; Conservative 39; Mismatches 94; Indels 102; Gaps 14;
QY 29 CENFPKMKIKLPFGTITIOFFYPQASFGISANTILLFHIFTFV-FSHRSKSIDMIISHL 87
Db 4 CLUVKDKVIRHAFAATIIIRRRVSVFIILGLIATMTPTFTKVFVFFQRCLSIMRFYSSL 63
QY 88 S---LIHLLLTQAI-----LVSLDFGSONTODDLRYKVIVFLNKMVRLGSLICTPCL 138
Db 64 PTFILIEIAMLFFMSVTCFLRCLSIIRFYSIST-----FILIDFVMPFFTLFTYFL 115
QY 139 --LSVLQAIISPSISLAKLHPSASHILGFFLFSWVLNMFIVFCCTLRPLPPVKRQOS 196
Db 116 RCLSIMR-----FSFS-----LITRIDFVMPFFMSV--TCFLRCLSIIRFYS 157
QY 197 SVCHTALFLFAHELHPQETVHTNDFEGCHLYRVHGPLKRLHGDYFIQ---TIRGYLSAF 253
Db 116 RCLSIMR-----FSFS-----LITRIDFVMPFFMSV--TCFLRCLSIIRFYS 157
QY 197 SVCHTALFLFAHELHPQETVHTNDFEGCHLYRVHGPLKRLHGDYFIQ---TIRGYLSAF 253

Db 158 SI---STFILIDFVMPFFTLF-----TYFLRCLSIIRFYSIS 192
QY 254 TOPACPRVSPVKRASQAILLVSVFVYVWDFT-----FSFSGVGTWINDSLVW 303
Db 193 T-----FILIDFVMPFFTLFTYFLRCLSIIMRFSF-----SLTTF 226
QY 304 LQV 306
Db 227 IRI 229

RESULT 14
US-08-990-571-32
; Sequence 32, Application US/08990571
; Patent No. 6214971
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G. et al.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B.
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990,571
; FILING DATE: 11-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-990-571-32

Query Match 4.6%; Score 85; DB 4; Length 245;
Best Local Similarity 22.4%; Pred. No. 0.39;
Matches 68; Conservative 39; Mismatches 94; Indels 102; Gaps 14;
QY 29 CENFPKMKIKLPFGTITIOFFYPQASFGISANTILLFHIFTFV-FSHRSKSIDMIISHL 87
Db 4 CLUVKDKVIRHAFAATIIIRRRVSVFIILGLIATMTPTFTKVFVFFQRCLSIMRFYSSL 63
QY 88 S---LIHLLLTQAI-----LVSLDFGSONTODDLRYKVIVFLNKMVRLGSLICTPCL 138
Db 64 PTFILIEIAMLFFMSVTCFLRCLSIIRFYSIST-----FILIDFVMPFFTLFTYFL 115
QY 139 --LSVLQAIISPSISLAKLHPSASHILGFFLFSWVLNMFIVFCCTLRPLPPVKRQOS 196
Db 116 RCLSIMR-----FSFS-----LITRIDFVMPFFMSV--TCFLRCLSIIRFYS 157
QY 197 SVCHTALFLFAHELHPQETVHTNDFEGCHLYRVHGPLKRLHGDYFIQ---TIRGYLSAF 253
Db 158 SI---STFILIDFVMPFFTLF-----TYFLRCLSIIRFYSIS 192
QY 254 TOPACPRVSPVKRASQAILLVSVFVYVWDFT-----FSFSGVGTWINDSLVW 303

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 18, 2002, 10:48:59 ; Search time 20.9 Seconds
(without alignments)
1641.335 Million cell updates/sec

Title: US-09-728-309-3
Perfect score: 1855
Sequence: 1 MLKLVIIENNAEIMFLSLD.....YLSPPKMLKFNQCGSTKK 357

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	365.5	19.7	310	2 I61748	phomone receptor
2	329.5	17.8	311	2 A57223	phomone receptor
3	312.5	16.8	321	2 I61749	phomone receptor
4	275.5	14.9	273	2 I61747	phomone receptor
5	269.5	14.5	335	2 I61746	phomone receptor
6	109	5.9	376	2 I50102	Phel3 bombesin rec
7	108.5	5.8	571	2 S50331	NADH dehydrogenase
8	106	5.7	409	2 S26033	NADH dehydrogenase
9	105.5	5.7	355	2 T11602	NADH dehydrogenase
10	104	5.6	361	2 F69269	hypothetical prote
11	102	5.5	713	2 A81317	probable integral
12	101.5	5.5	361	2 D72384	conserved hypothet
13	101	5.4	262	2 E97151	ABC transported MD
14	101	5.4	662	2 E97973	hypothetical prote
15	100	5.4	331	2 T21156	hypothetical prote
16	100	5.4	487	2 F72126	ct339 hypothetical
17	100	5.4	487	2 A81545	conserved hypothet
18	100	5.4	487	2 E86495	CT339 hypothetical
19	99.5	5.4	394	2 S48522	cell division cont
20	99	5.3	355	2 T14099	NADH dehydrogenase
21	99	5.3	786	2 S67060	probable membrane
22	98.5	5.3	496	2 E82940	hypothetical prote
23	98	5.3	355	2 T14123	NADH dehydrogenase
24	98	5.3	527	2 S53835	NADH dehydrogenase
25	98	5.3	662	2 E95105	ABC transporter, p
26	97.5	5.3	451	2 JE0166	nitric-oxide reduc
27	97.5	5.3	689	2 T32076	hypothetical prote
28	97	5.2	482	2 S63659	NADH dehydrogenase
29	97	5.2	528	2 S26025	NADH dehydrogenase

30	96.5	5.2	316	2 H86665	ferrichrome ABC tr
31	96.5	5.2	409	2 S26021	NADH dehydrogenase
32	96	5.2	465	2 T16618	hypothetical prote
33	95.5	5.1	355	2 T12114	NADH dehydrogenase
34	95.5	5.1	485	2 E88114	protein F53c3.9 [1
35	95	5.1	328	2 T28740	hypothetical prote
36	95	5.1	349	2 T33203	hypothetical prote
37	95	5.1	355	2 T13987	NADH dehydrogenase
38	95	5.1	355	2 T13982	NADH dehydrogenase
39	95	5.1	355	2 T14111	NADH dehydrogenase
40	94.5	5.1	302	2 B81696	4-hydroxybenzoate
41	94.5	5.1	355	2 T13833	NADH dehydrogenase
42	94.5	5.1	355	2 T12119	NADH dehydrogenase
43	93.5	5.0	355	2 T12112	NADH dehydrogenase
44	93.5	5.0	355	2 T13989	NADH dehydrogenase
45	93.5	5.0	355	2 T14030	NADH dehydrogenase

ALIGNMENTS

RESULT 1
I61748
phomone receptor VN6 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 29-Sep-1999
C:Accession: I61748
R:Dulac, C.; Axel, R.
Cell 83, 195-206, 1995
A:Title: A novel family of genes encoding putative phomone receptors in mammals.
A:Reference number: A57223; MUID:96028094
A:Accession: I61748
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-310 <RES>
A:Cross-references: EMBL:U36898; NID:g1055253; PIDN:AAC52287.1; PID:g1055254
C:Superfamily: phomone receptor VN3t

Query Match 19.7%; Score 365.5; DB 2; Length 310;
Best Local Similarity 32.1%; Pred. No. 8.5e-24;
Matches 104; Conservative 54; Mismatches 111; Indels 55; Gaps 10;

QY	39	LPQGITITQIFFYPOASPGISANTILLFLHIFTVFVSHRSKSIDMIISHLIHLILFTQ	98
DB	7	LYGVVDKQAIFFSEVWIGISFNSILFLHFQFLERRRITDLIISLLALHLGLMTVM	66
QY	99	AALVSLDFGSGONTQDRLRVKIVFLNKVMRGLSICITPCLLSVLQAI-ISPSIESLAKLK	157
DB	67	GFR-AVDIFASQNVWMDIKCKSLAHLRLGLSLCATCLLSITFOAITLSPRSSCLAKFK	125
QY	158	HPSASHITLGFFLFSWLVNMFIVFCCTLRPLPPVKRQSSVCHTALFLFAHELHPQETVF	217
DB	126	YKSTQHSLSILVLAWFYMSCGTHYSFTI-----VADYNFSSRSLSIF	167
QY	218	HTNDFEGCHLYRVHGPRLKRLHGD-YFTQTI-----RGYLSAF-----	253
DB	168	VT---ESCIIIL---PMDYITRDLFFILGIFRDVFSFGLMALSSGYMVALCRRHQAQH	220
QY	254	--TOPACPRYSPVKRASQAIIILVSV-FVFTYWDVDFTSFGGVTWINDSLVWLQIVAN	310
DB	221	LHRTSLSPKASPEORATVRIILLMSFVLMYCLDCTISAS-RLMHNGEPIHHSIQMMVSN	279
QY	311	SYAAISPLMIYADNQIFKTIQML	334
DB	280	SYATLSPLLLIVTENRISREKSL	303

RESULT 2
A57223
phomone receptor VN3 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999

[illegible]

A:Cross-references: EMBL:U36896; NID:g1055249; PIDN:AAC52285.1; PID:g1055250
C:Superfamily: pheromone receptor VN3t

Query Match 14.5%; Score 269.5; DB 2; Length 335;
Best Local Similarity 25.7%; Pred. No. 1.5e-15;
Matches 90; Conservative 69; Mismatches 132; Indels 59; Gaps 10;

QY 17 SLDLFFSTO-----ILCFNFPKMIK---LPGFITIOIEFFPOASFGISANTILLFHI 68
DB 2 SIEIILCHTQLWVVFVAFYPHTMKNKDLHVDTINKITWFSEVSGIILANSILFPGHL 61
QY 69 FTFVSHRSKSIDMIISHLIHLFTQAILVSLDFGQNTQDDLRYKVIIVFLKVM 128
DB 62 CMLIGENKRPRIHLIYASLSLTQMLLITMG-LIAADMTISQGIWDSTSCSLIYLRHS 120
QY 129 RGLSICTPCLLSVLQAI-ISPSIFSLAKLHPSASHILGFFLSWVLM----- 176
DB 121 RGFTLSAACLNLNVFMITLSKKKSCLTFKHNSPHHISGAPFLLLCVLYMCFSSHILISII 180
QY 177 -----FIGVIFCCTLRPLPVKRGQSSVCHTALFLFAHELHPQETVF-----HTNDF 222
DB 181 ATPNLITSDNFMVTKSCSF-LP-----MCYSRTSMFSTTIAVRAFFIGLMALSSGY 231
QY 223 EGCILYRVHGPLRLKHGDYFIQITIRGYLSAFTQACPRVSPVKRASQAIIILLVSVFTY 282
DB 232 LVAFRLWRHRRKQAOHLHS-----TGLSS-----KSSPEQRATETILLMSFFVLY 276
QY 283 VDTFFSGGVTTWINDSLVQLVIVANSYAATISPLMLIYADNOIFKTLQ 332
DB 277 ILENVVFYSRMKFGDSTFYCVQIIVSHSYATVSSFFVIFTERKMTKILR 326

RESULT 6
I50102
Phe13 bombesin receptor - Bombina orientalis
C:Species: Bombina orientalis
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 04-Sep-1998
C:Accession: I50102
R:Nagalla, S.R.; Barry, B.J.; Creswick, K.C.; Eden, P.; Taylor, J.T.; Spindel, E.R.
Proc. Natl. Acad. Sci. U.S.A. 92, 6205-6209, 1995
A:Title: Cloning of a receptor for amphibian [Phe13]bombesin distinct from the receptor
A:Reference number: I50100; MUID:95320240
A:Accession: I50102
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-376 <NAG>
A:Cross-references: GB:L39358; NID:g903700; PID:g903701
C:Genetics:
A:Gene: B84
C:Superfamily: endothelin receptor B

Query Match 5.9%; Score 109; DB 2; Length 376;
Best Local Similarity 21.1%; Pred. No. 0.084;
Matches 68; Conservative 49; Mismatches 120; Indels 86; Gaps 15;

QY 54 SFGISANTILLFHIFTFVFSHRKSIDMIISHLIHLFT-----QAILVSLDFF 107
DB 58 SVGILGNTILI--KVF-FKIKSMQTVPNIFTSLAFGDLILLLTFCVPDASRYIVDTWTF 114
QY 108 GSQNTQDDLRYKVIVFLNKNVWRLGSLICTPCLLSV--LQAIISPSIFSLAKLHPSASHIL 165
DB 115 GRACG-----KIISFTQLTSVGVSVFTLAVSLTDTRYRAIVKP-----LQLQTSDAVL 161
QY 166 ---GFFLFSLVNLNMFIVGICCTLRPLPVKRGQSSVCHTALFLFAHELHPQETVFHTNDF 222
DB 162 KTCGKAVCVMIISMLL-----AAPEAVFSDLYEFGSSE---KNTTF 199
QY 223 EGCILYRV-----HG-----PLKRLHGDYFIQITIRGYLSAFTQ-----AC 258
DB 200 ECAPYVPSEKILQETHSLICFLVYIVPLSIISAYFLIAKTLKSTFNPAPAEHTAR 259

QY 259 PRVSPVKRASQAIIILLVSVFTYVW-----DFTFSGGVTTWINDSLVQLVQIVA 309
DB 260 KQIESRRKRVAKTVLVLVALFAVCWLPNHNMLYLYRSFTYHSAVNSSAFHLSATIFARVAL 319
QY 310 NSYAAISPLMLIYAD---NOIFK 329
DB 320 RN-SCVNPEFALYWLRSFRQHF 341

RESULT 7
S50331
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Katharina tunicata mitochondri
C:Species: mitochondrion Katharina tunicata
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-Dec-1999
C:Accession: S50331
R:Boore, J.L.; Brown, W.M.
Genetics 138, 423-443, 1994
A:Title: Complete DNA sequence of the mitochondrial genome of the black chiton, Katha
A:Reference number: S50327; MUID:95129806
A:Accession: S50331
A:Molecule type: DNA
A:Residues: 1-571 <BOO>
A:Cross-references: EMBL:U09810; NID:g557273; PIDN:AAC48366.1; PID:g557278
C:Genetics:
A:Gene: ND5
A:Genome: mitochondrion
A:Genetic code: SGC4
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation

Query Match 5.8%; Score 108.5; DB 2; Length 571;
Best Local Similarity 21.5%; Pred. No. 0.14;
Matches 76; Conservative 53; Mismatches 146; Indels 79; Gaps 15;

QY 34 SKMILKPGFITIQIIFYPOASFGISANTILLFHIFTFVFSHRKSID-----MIISHL 87
DB 220 STIVTAGVFLIQL--FYPLFSLNSHLFSTSLMIISMTVMVAGISANFETDLKIIALSTL 277
QY 88 SLIHLILFTQAILVSLDFGQNTQDDLRYKVIIVL-----NKMVRLGLSTCTP 136
DB 278 SOLGVMLSVSLGLFSLALF--HLFTHAMFKALLFLCAGNIHSHNNQDIRKMSHLW- 333
QY 137 CLLSVLOAIISPSIFSLAKLHPSASHILGF-----FLFSVNLNMFIV 180
DB 334 -----IQMFESTCFNIANLALACGFPFMAGFYKDVIIEMMMFNQNNFLIS--LNMELAT 386
QY 181 IFCTLRPLPVKRGQSSVCHTALFLFAHELHPQETVFHTNDFECCHLYRVHGPLKRL--- 237
DB 387 ILTA-----SYSARLSLUFIFWSKM-TOSSVSVSTSDEN-----SIAPUTMLAAG 430
QY 238 ---HGDFYFIQITIRGYLSAFTQACPRVSPVKRASQAIIILLVSVFTYVWDTFSGSVT 294
DB 431 AITSGAVFSLI-----LTPSTLPLISMNMKLMATATLTGRAITYKL-FKLELKISHS 483
QY 295 WINDSL--VNLQVIVANSYAATISPLMLIYADNOIFKTLQMLNFKYLPKMLK 347
DB 484 IINFFFNMFMINLSNNMTSKSL---YTGLYTFKSLDLGWMETIGGLEGLNK 534

RESULT 8
S26033
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Caenorhabditis elegans mitoch
C:Species: mitochondrion Caenorhabditis elegans
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 07-Dec-1999
C:Accession: S26033; S25806
R:Okimoto, R.; Macfarlane, J.L.; Clary, D.O.; Wolstenholme, D.R.
Genetics 130, 471-498, 1992
A:Title: The mitochondrial genomes of two nematodes, Caenorhabditis elegans and Ascar
A:Reference number: S26014; MUID:92201635
A:Accession: S26033
A:Molecule type: DNA
A:Residues: 1-409 <OKI>

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5.
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation

```
Query Match          5.7%; Score 105.5; DB 2; Length 355;
Best Local Similarity 20.4%; Pred. No. 0.16;
Matches 75; Conserved 59; Mismatches 129; Indels 105; Gaps 19;

QY  4  LVIENMAEIMLSLDLLLFSDILCFNFPFSKMILKPGFTIQIFVYPOASFGI-----57
DB 17  MILNENCLFELLISSUTFMAG-LGANFEEDLKIIAISTL-----SGLGLMMSILS 68
QY 58  SANTILLLFHIFTVFVSHRSKSIDMIISHUSLIHILLFTQAILVSLDFFGSONTODDLR 117
DB 69  MGNVYKLAFFHLTH-----ALFKALLFCAGVII---HNLKDTQD---105
QY 118  YKIVVFNKYWRGLSICTPCLLSVLQAILSPFSFSAKLKHPSPASHILGFFLESWVL---174
DB 106  ---IRFNGMLVMIMPLTICM-----NISNLALCGMP---FLAGFYSKDLILEVY 149
QY 175  ----NNFIVGIFECCTLRLPPVKRGSSVCHTALFLFAHELHPQETVFHTNDEGCHLYR 229
DB 150  SMPDINIFILFISVGL-----TVCTYRLEY-----YTI--TSDF---NFYS 189
QY 230  VHG-----PLKRLHGDIYFIQTIRGYSLAFQACPRVSPVKRRAQAILLVSPFYETW 282
DB 190  LHSLNDEGWIMKSLMLLMFVIFSGSMLMWLIFPTPYMCLPIELKMLALLVSPFGA-W 248
QY 283  VDTTF-SFSGGVWINDSL-----LVWLOVIVANSYAAISPLMLIYADNQIFKYL 331
DB 249  LGYEMASFS--VSWVSNLSLGFNYFFGFGPMFLPTISTFSVNVYVPLMLSY---NLFSKF 303
QY 332  QMLWFKYL 339
DB 304  DQGNWYL 311

RESULT 10
F69269
hypothetical protein AF0158 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: F69269
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
.: Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A:Reference number: A69250; MUID: 98049343
A:Accession: F69269
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-361 <KLE>
A:Cross-references: GB:AE001095; GB:AE000782; NID:g2689418; PIDN:AAB91076.1; PID:g265
```

Query Match 5.6%; Score 104; DB 2; Length 361;
Best Local Similarity 20.2%; Pred. No. 0.22;
Matches 71; Conserved 56; Mismatches 129; Indels 96; Gaps 15;

QY 12 EIMLFSDLLLFDILCFNFPFSKMILKPGFI--TQIFVYPOASFGISANTILLFHIF 69
DB 22 EPVAFSYALLISGADLLL---AALALLSYLRRRAIPMFLI---LGLSFFSVLLIGPLA 74
QY 70 TFFVFSRKSIDMIISHLSL-----IHIL-----LLFTQAILVSLDFFGSONTQD 114
DB 75 DLALPHRATEI-LTRPHLASTEMHPGISVVALYGLLWPLTFVIFALLFYSPYPMHK 133
QY 115 DLRKYTVIVFNKYWRGLSICTPCLLSVLQAIISPSIFSFAKLKHPSPASHILGFFLESW--172
DB 134 GGTFPSLFSFGVKSSEYERLKPA-MKVLAAILVP--LSALWTIYIPGMLFFSQTQWIAVKN 190

	Query Match	5.5%	Score 101.5;	DB 2;	Length 361;
	Best Local Similarity	20.2%;	Pred. No. 0.35;		
	Matches 68;	Conservative 50;	Mismatches 108;	Indels 111;	Gaps
Oy	48 FFYQASFGISANTILLFHIFTFVSHRSKSIDM----	IISHLSLHILLLFTOAILVSL 104			
	: : : : : : : : : : :				
Db	24 FLFP-----	ILLYLILYSIFGDMSSNMNLRLGLVGESSLVESFHLPS--GI 69			
Oy	105 DFFGSQNTQDDLRYKVIVFLKNVRGLSICITPCLLSVLQAIISP-	-----IFSLA 154			
	: : : : : : : : : : : :				
Db	70 EIIHQTPERDLKYGKI-----DVC-----	VILPENFDNLTKAIIFSKT 109			
Oy	155 KLKHPSSASHIL-----	GFFLSWVLNMFIGVICCTLRLPVPKRGQ----- 195			
	: : : : : : : : : :				
Db	110 KISVPEVETILYAPERQESIVLANNVSN----	ILESDLQLRAKVEVEVEYRKKEQAIN 165			
Oy	196 -----	SSVCHTALFLFAHELHPQETVFHTNDFEGCHLYRVHGPKLRKHGDYFIQ 244			
	: : : : : : : : : :				
Db	166 YSHILPAVLVLGVMSGLFTVPVQL-----	ALYREGIUKRI-----LAS 206			
Oy	245 TIRGYLSAFTOPACRPSPVKRASQAILLIVSFVTYVVDFTFSGGV-	TWINDS 299			
	: : : : : : : : : :				
Db	207 PLRHHHFVSFVVC-GLFAMSIITSASTLTARFYVGEFFPINVSFLGCVLLSEFLTFSV	265			
Oy	300 LLVWLQIVANSAAISPLMLIYADNOIEKTLQMLWF 336				
	: : : : : : : : : :				
Db	266 LLM---VSLFTFALSAAISOVF--NOVMFLGGFYF 297				
RESULT	13				
E97151					
ABC	transported MDR-type, permease component CAC2041 [imported] - Clostrid				

C:Accession: E97151
C:Accession: E97151
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gil-
R: Dally, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J: Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing
A:Reference number: A96500; MUID:21359325; PMID:21359325
A:Accession: E97151
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK80000.1; PID:q15025026; GSPDB:GN0
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2041

Query match	5.4%	Score 101	DB 2	Length 262
Best Local Similarity	23.8%	Pred. No. 0.28		
Matches 53	Conservative 29	Mismatches 73	Indels 68	Gaps

Qy	7	IENNAEIMFLSDDLLESTDI	LCFNFPKMKIKLPGFIT	IOIFF-YPOASFGISANTILL	65
		: : : :	: : :	: : :	
Db	52	VENNSFIITFTPICLISAY	IFSRFSYK-----TSOLFCYPOSRTKIFISKLLTI		103
		: : :	: : :	: : :	
Qy	66	FHIITFVF-----SHRS	SIDMIHSLIHILLFTQ-----AIVSLD	105	
		: : :	: : :	: : :	
Db	104	MVLFAFLLELLVITILLG	PMCPHEVLTQSIIYHFKL-YLYVLFECTISPICILISLL	162	
		: : :	: : :	: : :	
Qy	106	FFGSQNTQDDLRKYKIV	-----FLNKVMRG-----LSICTPCLLSVLQA	144	
		: : :	: : :	: : :	
Db	163	F---KNTWTPLIYGILIS	VCNLFISTFVYVKNANGGLMKNIIFNIPLFYNVPIYSCFA	219	
		: : :	: : :	: : :	
Qy	145	IISPSIF-----SLAKL	KHPNSHILGFFLFSWLNMEIVGFC	183	
		: : :	: : :	: : :	
Db	220	NGKAIFIESSILSHH-----	IFICILTFTIANLFC	252	
		: : :	: : :	: : :	

E97973
hypothetical protein ABC-MSD [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: E97973
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
y, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: E97973
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-662 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAK99617.1; PID:g15458413; GSPDB:GN00174
C:Genetics:
A:Gene: ABC-MSD

Query Match 5.4%; Score 101; DB 2; Length 662;
Best Local Similarity 21.6%; Pred. No. 0.72;
Matches 74; Conservative 57; Mismatches 116; Indels 96; Gaps 15;
QY 6 IENNAEIMFLSLLDILCF---NFPKMKILPFTIIFYPQASFGISANTI 62
DB 13 LIKRNKLYPFALVAVTFLVFLYLTNPKIAIRGGTTIQ---ATLFGFMVVT 68
QY 63 ---LFLHFTFVFSHRKSIDMILSHLSLHILLFTQAILVSLDFFGSONTQDDRLYK 119
DB 69 ASAIIVLYANSVMKRSKELG-----YGMGLGKRLHLSMTF-----KE 109
QY 120 VIVF-LNKVMRGLSICTCLLSVLQAIISPSIFSLAKLHPQSASHILGFFLFSWVLN--- 175
DB 110 LVVFGILVAGIGIG-----ALFDKLIFALLKMLK-----VELVATFQTKVITVIV 160
QY 176 ----MFIGVIFCCTLRP-----PVKRGSSVCHTALFLFAHELHPQETVHTNDFEGCH 226
DB 161 VFGLIFGLMFLNALRIAMNALQLSREKASGEKKGREL-----PLQTL----- 205
QY 227 LYRVHGPLKRLHGDYFIOTIRGYLSAFTQPCPRVSPVKRASAAILLLVSVFVTVVDFT 286
DB 206 ----GTSISLGYLYALTVDPLTALT-----TFFIALLVIFGTYLL---- 244
QY 287 FSGSGVTWINDSLVWLQVIVANSYAAISPLMLIYADNQIFK 329
DB 245 --FNAGIT-----VFLQILKKNKKYYPNNLISVSNLIPIR 278

RESULT 15
T21156
hypothetical protein F20E11.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 26-May-2000
C:Accession: T21156
R:Alusough, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19383
A:Accession: T21156
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-331 <WIL>
A:Cross-references: EMBL:Z81508; PIDN:CAB04142.1; GSPDB:GN000023; CESP:F20E11.1
A:Experimental source: clone F20E11
C:Genetics:
A:Gene: CESP:F20E11.1
A:Map position: 5
A:Introns: 58/3; 94/2; 191/3; 245/3
C:Superfamily: Caenorhabditis elegans hypothetical protein K02H11.4

Query Match 5.4%; Score 100; DB 2; Length 331;

Best Local Similarity 22.3%; Pred. No. 0.44;
Matches 71; Conservative 41; Mismatches 97; Indels 110; Gaps 15;
QY 47 IFFYPQASFGISANTILLHFLHFTFVFSHRKSIDMI-----ISHLSLIHILL 95
DB 26 IFLYFFAF-IIIYITLLPFFYVFTNKLHNRDNEMFIYPTAHFCOMKVSYLVFVSACFN 84
QY 96 F-----TQAILVSLDFFGSON-----TQDDRLYKVIIVL-----NK 126
DB 85 FILAVSVNSLQTNRVLIALLFPALFNVLVIITQ---AFHVLIFVLAVERFFIYFFPSSEK 141
QY 127 VMRGLSICTPCLLSVLQA-IISPSIFSLAKL-----HPSASHILGFFLFSWVLN---M 176
DB 142 IFKLSIKIKHLYFYIIAKDFSCFVTVMKRNAAFQTVFQNLTWYDFELMTWCINDIIV 201
QY 177 FIGV-----IFCCTLRLPVVKRGSSVCHTALFLFAHELHPQETVHTNDFEGCHLYRVH 231
DB 202 FISALLYPIFISVRKLVNLSQVQRKPHNYIFL-----QTMVVF----- 242
QY 232 GPLKRLHGDYFIOTIRGYLSAFTQPCPRVSPVKRASAAILLLVSVFVTVVDFTFSFG 291
DB 243 --LKSIIHILYIMTSHG-----GNFSP-----LFTNMIIVT----- 272
QY 292 GVTWINDSLVWLQVIVAN 310
DB 273 --DFITTPLIIVQMSYLSCN 289

Search completed: September 18, 2002, 10:53:00
Job time: 241 sec

```

FT DOMAIN 103 120 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 121 142 3 (POTENTIAL).
FT DOMAIN 143 162 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 163 183 4 (POTENTIAL).
FT DOMAIN 184 219 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 220 240 5 (POTENTIAL).
FT DOMAIN 241 271 6 (POTENTIAL).
FT TRANSMEM 272 292 6 (POTENTIAL).
FT DOMAIN 293 312 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 313 332 7 (POTENTIAL).
FT DOMAIN 333 392 9 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 9 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 119 202 BY SIMILARITY.
FT LIPID 346 346 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 392 AA; 43461 MW; 497C620209F1DBEB CRC64;

Query Match 6.0%; Score 112; DB 1; Length 392;
Best Local Similarity 21.7%; Pred. No. 0.11;
Matches 70; Conservative 48; Mismatches 119; Indels 86; Gaps 15;

QY 54 SFGISANTILLFIHFTVFVSHRSKSIDMIISHLISLHILLFT-----QAILVSLDF 107
DB 58 SVGILGNTILI--KVF-FKIKSMQTVNFIITSLAPGDELLLTCTVPVDASRYIVDTWTF 114

QY 108 GSQNTODDLRYKIVFLNKMVRGLSCTPCLLSV--LQAIISPSIFSLAKLKHPSASHIL 165
DB 115 GRAGC-----KIISFIQLTSVGVSVFTLTLSADRYAIRVXP-----LQLOTSDAVL 161

QY 166 ---GFFLFWNLNMFIVFCCTLRLLPPVKRGSSVCHTALFLFAHELHPQETVFTNDF 222
DB 162 KTCGKAVCVMIISMLL-----AAPEAVFSDLYEFGSSE---KNTTF 199

QY 223 EGCHLYRV-----HG-----PLKRLHGDFYFIQTIRGYL-SAFTQACPR-----VSPVKRASQAILL 273
DB 200 ECAPVPSEKILQETHSLICFLVFIIVPLSISAYIFLIATLYKSTNMPAEETHAR 259

QY 259 PRVSPVKRASQAILLVSVFTYVW-----DFTSFSGGVTVINDSLVLWLQVIVA 309
DB 260 KQIESRKRAKTVLVLVALFAVCWLPNHLMLYLRSTYHSYAVNSAFHLSATIFARVL-A 318

QY 310 NSYAAISPLMLIYAD---NQIEK 329
DB 319 FNSNCVNPFALYWLSRSFRQHEK 341

RESULT 2
NU4M_CAEEL STANDARD; PRT; 409 AA.
AC P24892;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN ND4.
OS Caenorhabditis elegans.
OG Mitochondrion.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=92201635; PubMed=1551572;
RA Okimoto R., Macfarlane J.L., Clary D.O., Wolstenholme D.R.;
RT "The mitochondrial genomes of two nematodes, Caenorhabditis elegans
and Ascaris suum.";
RL Genetics 130:471-498(1992).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X54252; CAA38158.1; -.
DR PIR; S26033; S26033.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 409 AA; 47206 MW; AFBFE452A5814F7C CRC64;

Query Match 5.7%; Score 106; DB 1; Length 409;
Best Local Similarity 20.4%; Pred. No. 0.33;
Matches 80; Conservative 73; Mismatches 148; Indels 92; Gaps 20;

QY 1 MLKLVIIENNAEIMFLSLLFLSDDILC--FNFSKMIKLPGF-----ITIQIFYYP 51
DB 57 ILGIIVISEKNNNLLILSLVLF---ICIIFFIPSNMMMLYMFELSMFPILVLMILGYG 112

QY 52 QASFGISANTILL---PHIETFVFSHRKSIDMIISHLISLHILLFTQAILVSLDFRG 108
DB 113 SOIEKINSYYLNFYAAFCSPFPLVFYKSNLLVFTYTNFV---ISWEMFFILSLSPM- 168

QY 109 SONTODDLRYKIVFLNKMVRGLSCTPCLLSVLQAIISPSIFSLAKLKHPSASHILGFF 168
DB 169 -----MKFP-IYFLHLWLPKRAHVEAPTASMLLAGL-----LLKLGTAGFLRLGSL 214

QY 169 LF-----SWVLNMFIVI---FCTLRLL-PPVKRGSSVCHTALFLFAHELHPQETVFTN 219
DB 215 SFVHNWNLIIAFLGMLGSCFQSDSKALAAVSVTHMSFLLS-----LVFITM 267

QY 220 -NDFECCHLYRVHGLKRLHGDYFIQTIRGYL-SAFTQACPR-----VSPVKRASQAILL 273
DB 268 SKKISSVMLMLAHG-----YTSLMFYLIGEYHTSGSRMIYFSSFSSSMIMGI 318

QY 274 LVSFVFTYVWDETFSGFSG---GVTWINDSLVLWLQVIVANSYAAISPLMLIYADNIFKT 330
DB 319 LFSVVF-----LSNSGVPPSLSLSEFLVISNSMLSKS---MFMVFIYFVVVSFYYS 368

QY 331 LQML-----WFKYLSPPKMLKEN 349
DB 369 LFLITSSLMKGYNHFNENTWNVGFSAFLVLMYNN 401

RESULT 3
CC91_YEAST STANDARD; PRT; 394 AA.
AC P41733;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cell division control protein 91.
GN CDC91 OR YLR459W OR L9122.2.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Bi E., Pringle J.R.;
RL Submitted (XXX-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favell A., Fulton L., Gattung S., Greco T., Kirsten J.,
RA Kuceba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,

```


DE NADH-ubiquinone oxidoreductase chain 2 (EC 1.6.5.3).

GN ND2 OR NAD2.

OS Acanthamoeba castellanii (Amoeba).

OG Mitochondrion.

OC Eukaryota; Acanthamoebidae; Acanthamoeba.

OX NCBI_TaxID=5755;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 30010 / NEFF;

RX MEDLINE=9514725; PubMed=7844823;

RA Burger G., Plante I., Loneragan K.M., Gray M.W.;

RT "The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba

castellanii: complete sequence, gene content and genome

organization.";

RL J. Mol. Biol. 245:522-537(1995).

CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial

inner membrane.

CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; U12386; AAD11827.1; .

DR InterPro: IPR001750; Oxidored.q1.

DR Pfam: PF00361; oxidored.q1.

KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.

SQ SEQUENCE 527 AA; 61407 MW; 417DABB22EE04F19 CRC64;

Query Match 5.3%; Score 98; DB 1; Length 527;

Best Local Similarity 22.1%; Pred. No. 1.7;

Matches 90; Conservative 53; Mismatches 138; Indels 126; Gaps 21;

QY 48 FFYPOASFGISANTLLLFHIFTFV---FSHRKSID---MIISLSLIHILLFTQAIL 101

DB 101 FYNDAAVVFKNILIGLIFTFAIKOYLSYFKYDYDFEILVLFISLSSLLILNSDL 160

QY 102 VSLDFP-----GQNTQDDLRKYKVIIFLNK---VMRGLSIC--TPCLLS 140

DB 161 ISLFFIIEQLSTFVLVASKQTSFSTESGKYPILGCFSSGIIIFGLISLYGTGLLS 220

QY 141 -----VLQAIISPSIFSLAKLHPSASHILGFFLSWVNNFIVGIFCC--- 184

DB 221 YTDLTFLSEVYVTFNILDSSFFSF-----SGFLIGLLLT-----VGFLKLGSA 267

QY 185 -TLRUPPVKRGQSSVCHTA-----LPLFAHELHPQETVFT 219

DB 268 FHMMPDVVEG-SPLLIATAYLSTLPKISLIFVIFKLYYVFEVFLFSQGLFTLTALESI 326

QY 220 NDFECCHLYRVHGPGLKRLHGDYFIOTIRGYLSAFTOPACRPSPVKRASQAILLVSVF 279

DB 327 LGSTAAYIQV--KUKRLM-TYSMTITNTGYLLGLS-----FGDISGYITIFYLISVIF 378

QY 280 TYWVDFTEFS-----SG-----GVTWINDSLVLMQIVANSYAASIPMLIY 322

DB 379 -IMIGLFCFLSLDRSSGLLVKRLNLSNLEVNPSLSFSIFILLF-SIAGIPPLLGFY 436

QY 323 ADNQIF---KTIQMLW-----FKYLSPPKMLKFNROCG 353

DB 437 SKFFLFLESLKMYKMYMTILFVFSWSVYFVIRLVKLMY-FNRRTG 482

RESULT 6

NU5M_ASCSU

ID NU5M_ASCSU

AC P24884;

DT 01-MAR-1992 (Rel. 21, Created)

STANDARD; PRT; 547 AA.

DT

DE

OS

OG

OC

OX

RN

RP

RC

RX

RA

RT

RL

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

DR

DR

DR

KW

SQ

Query Match

Best Local Similarity

Matches

QY 4 LVIENMAEI-----MLFSLDLLLFTST-----DILC 29

DB 84 LVFVGSMEFLIFSSGCFSMVSWDLLGISSFLVLFYNNWDSGAMTVLTNRLGDFL 143

QY 30 ENFPSKMLKPGFITIQIFFYPOASFGISANTLLLFHIFT---FVES---HRSKSIDM 82

DB 144 FVFFSSTI-----FSSYIFLSUSFFCWLSSMLLLASFTKSAQFPFSGMLPKRAMSAPT 196

QY 83 IISLSLIH-----ILLFTQAIL-----VSLDFGQSN--TQDDLRY 118

DB 197 PIS--SLVHSSLTLYTAGLVLMNFSEMIINLKDVIIMVGVTFMFFSSMAALVEEDLK- 253

QY 119 KVIPLNKVMRGLSTCTPCLLSVLQAIISPSIFSLAKLHPSASHILGFFLSWVNNFI 178

DB 254 KVALSTLSQMGFSMLTVGI-----GLSFVSFIHLLSHALFKSLFMQV 297

QY 179 GVIFCCTLRLEPPVKRCQSSVCHTALFAHELHPQETVFTHTNDFEGCHLYRVHGPLKRLH 238

DB 298 GYLHICSLGQDD-GRNYSNLGNVPYFI---QLQLLVTLFCL-----CGLVFSSGAVSK-- 346

QY 239 GDYFIQTI-----RGYLSAFTOPA-CPRVSPVKRASQ 269

DB 347 -DYILEFFFSNFFVVFACMEFFSVLTFGYSLRWLKGFFMFSRPFCSFSSVVMNLS 405

QY 270 AILLVSVFTYVDF 285

DB 406 LLLVLSIFFIWMNF 421

RESULT 7

NU4M_ASCSU

ID NU4M_ASCSU

AC P24880;

DT 01-MAR-1992 (Rel. 21, Created)

STANDARD; PRT; 409 AA.

Qy 278 VFTYVWDFTS-----FSGG-----VTWINDSLVWQVIVANSYAAI-----SPLMLIY 322
 Db 273 ITRTLRQLSPERAFSGKAALTLAFVQGLVLCWLPFFFIHLOMSITGSKSPGDLEE 332
 Qy 323 ADNQIFKTLQMLWFYLS---PPKL-----MLKFNRCQ 352
 Db 333 AVN-----WLAYSFAVNPSPYGLLRQIRDELVKFRCC 367

RESULT 9
 PSAA_EUGGR STANDARD; PRT; 751 AA.
 AC P19430;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Photosystem I P700 chlorophyll A apoprotein A1 (PsaA) (PSI-A).
 GN PSAA.
 OS Euglena gracilis.
 OG Chloroplast.
 OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
 OX NCBI_TaxID=3039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z;
 RX MEDLINE=88223484; PubMed=2836086;
 RA Cushman J.C., Hallick R.B., Price C.A.;
 RT "The two genes for the P700 chlorophyll a apoproteins on the Euglena
 gracilis chloroplast genome contain multiple introns.";
 RL Curr. Genet. 13:159-171(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z;
 RX MEDLINE=93347989; PubMed=8346031;
 RA Hallick R.B., Hong L., Drager R.G., Favreau M.R., Monfort A.,
 RA Orsat B., Spielmann A., Stutz E.;
 RT "Complete sequence of Euglena gracilis chloroplast DNA.";
 RL Nucleic Acids Res. 21:3537-3544(1993).
 RN [3]
 RP SEQUENCE OF 1-398 FROM N.A.
 RA Manzara T., Hu J.X., Price C.A., Hallick R.B.;
 RT "Characterization of the trnD, trnK, psaa locus of Euglena gracilis
 chloroplast DNA.";
 RL Plant Mol. Biol. 8:327-336(1987).
 CC -!- FUNCTION: PsaA and psaa bind P700, the primary electron donor of
 photosystem I (PSI), as well as the electron acceptors A0, A1, and
 FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin
 oxidoreductase.
 CC -!- COFACTOR: P700 is a chlorophyll A dimer, A0 is chlorophyll A, A1
 is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.
 CC -!- SUBUNIT: A psaa/B heterodimer binds the P700 chlorophyll special
 pair and subsequent electron acceptors. The PSI reaction center
 of higher plants and algae is composed of at least 11 subunits.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
 thylakoid membrane.
 CC -!- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; X70810; CAA50093.1; -
 DR EMBL; Z11874; CAA77910.1; -
 DR EMBL; M37526; AAA84451.1; ALT_SEQ.
 DR EMBL; M17309; AAA84232.1; -
 DR PIR; S26071; S26071.
 DR Mendel; 4408; EUGgr:psaa.1.
 DR InterPro; IPR001280; PsaA_psaB.

DR pfam; PF00223; psaa_psaB; 1.
 DR PRINTS; PR00257; PHOTOSPSAAB.
 DR PROSITE; PS00419; PHOTOSYSTEM_I_PSAAB; 1.
 KW Photosynthesis; Photosystem I; Electron transport; Chloroplast;
 KW Thylakoid; Transmembrane; Iron-sulfur; 4Fe-4S; Chlorophyll.
 FT TRANSMEM 73 96
 FT TRANSMEM 159 182
 FT TRANSMEM 198 222
 FT TRANSMEM 294 312
 FT TRANSMEM 349 372
 FT TRANSMEM 388 414
 FT TRANSMEM 436 458
 FT TRANSMEM 533 551
 FT TRANSMEM 591 612
 FT TRANSMEM 665 687
 FT TRANSMEM 725 745
 FT METAL 575 575
 FT METAL 584 584
 FT BINDING 676 676
 FT BINDING 684 684
 FT BINDING 692 692
 FT BINDING 693 693
 FT BINDING 718 718
 SQ SEQUENCE 751 AA; 83988 MW; C3F88ADF3AA2295C CRC64;

Query Match 5.0%; Score 93; DB 1; Length 751;
 Best Local Similarity 21.7%; Pred. No. 5.9;
 Matches 80; Conservative 52; Mismatches 148; Indels 88; Gaps 17;

Qy 2 LKLVIIENMAEIMFLSDDLFLSTDLFCNFPKMIKLPGFITI-----QIF----FYP 51
 Db 363 LSIIVAOHMYSMPPYPYIAIDYGTSLF---THYIWGGFCIVGAAHAAAFMVRDYP 419
 Qy 52 QASFGISANTILLFIHFFVESHRSKSIDMIISHLIHLILFTQAILVSLDFFSQ 110
 Db 420 ALNENLLDRVLL-----HR-----DAIISHLNWCIFLGLHSFLYIHNDLSAL 465
 Qy 111 NTQDDLRYKVIVFLNKMVRLSICTPCLLSVLQA---IISPSI-----FSLAK 155
 Db 466 GRPQDMFSDSAIQLOPVEAQWQIOTNYLAPTATFNLYSPTTPVWGGDVVISGKVAMP 525
 Qy 156 LKHPSAS-----HILGFFLFSWLNMFIVFCCTLRLLPVRK-----GOSSVC- 199
 Db 526 IKLGTDADFLVHHIAFTIHVTVILLKGLFSRSSRLIPDKASLFRFPDGGRGTCQ 585
 Qy 200 -----HTALELFAHELHPQETVEH-----TNDFECHLYRVHGLPLKRLHGDYFIO---TI 246
 Db 586 VSAMDHVELGLFWYNSISVAIFHFWSKMQSDVWGTVL---ANKVSHITGNFSGSLTI 642
 Qy 247 RGYLSAFTQACPRV-----SPVKRASQAIIILLVSFVFTYWWDFTFPSG-----GVT 294
 Db 643 NGWLRDFLWAQSSQVTSYSGSPLS-AYGLMFLGAHFVWAFILMFLFSGRGYQWELIESIV 701
 Qy 295 WINDSLV 302
 Db 702 WAHNLKV 709

RESULT 10
 CCAB_DISOM STANDARD; PRT; 2326 AA.
 ID CCAB_DISOM
 AC P56698;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Probable voltage-dependent N-type calcium channel alpha-1B subunit
 DE (DOE-4).
 OS Discopyge ommata (Electric ray).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Squala; Hypnosquala; Pristioraja; Batoidea;
 OC Torpediniformes; Narcinoidei; Narcinidae; Discopyge.

OX NCBI_TaxID=7785;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Electric lobe;
 RX MEDLINE=93248175; PubMed=7683405;
 RA Horne W.A., Ellnor P.T., Inman I., Zhou M., Tsien R.W., Schwarz T.L.;
 RT "Molecular diversity of Ca²⁺ channel alpha 1 subunits from the marine
 ray *Discopyge omata*";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:3787-3791(1993).
 CC -1- FUNCTION: THE ISOFORM ALPHA-1B GIVES RISE TO N-TYPE CALCIUM
 CC CURRENTS. N-TYPE CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE
 CC ACTIVATED" (HVA) GROUP (BY SIMILARITY).
 CC -1- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
 CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
 CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
 CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
 CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
 CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
 CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS ARE PRODUCED BY
 CC ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSION IS HIGHER IN THE ELECTRIC LOBE THAN
 CC IN THE FOREBRAIN.
 CC -1- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
 CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
 CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
 CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
 CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
 CC -1- PTM: PHOSPHORYLATED IN VITRO BY CAM-KINASE II, CAPK, PKC AND CGPK
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L12532; -; NOT_ANNOTATED_CDS.
 DR InterPro: IPR002077; Ca_channel.
 DR InterPro: IPR002111; Cat_channel_TpL.
 DR InterPro: IPR000636; Cation_chan_non_lig.
 DR InterPro: IPR001682; Channel_pore_Ca_Na.
 DR Pfam: PF00520; Ion_trans; 4.
 DR PRINTS; PR00167; CACHANNEL.
 DR Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
 KW Calcium channel; Glycoprotein; Repeat; Multigene family;
 KW Calcium-binding; Phosphorylation; Alternative splicing.
 FT REPEAT 75 351 I.
 FT REPEAT 458 702 II.
 FT REPEAT 1134 1416 III.
 FT REPEAT 1453 1708 IV.
 FT DOMAIN 1 88 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 89 107 S1 OF REPEAT I (POTENTIAL).
 FT DOMAIN 108 125 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 126 145 S2 OF REPEAT I (POTENTIAL).
 FT DOMAIN 146 156 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 157 176 S3 OF REPEAT I (POTENTIAL).
 FT DOMAIN 177 180 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 181 199 S4 OF REPEAT I (POTENTIAL).
 FT DOMAIN 200 218 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 219 238 S5 OF REPEAT I (POTENTIAL).
 FT DOMAIN 239 323 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 324 348 S6 OF REPEAT I (POTENTIAL).
 FT DOMAIN 349 472 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 473 491 S1 OF REPEAT II (POTENTIAL).
 FT DOMAIN 492 506 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 507 526 S2 OF REPEAT II (POTENTIAL).

FT	DOMAIN	527	534	CYTOPLASMIC (POTENTIAL).
FT	TRANSEM	535	552	S3 OF REPEAT II (POTENTIAL).
FT	DOMAIN	553	563	EXTRACELLULAR (POTENTIAL).
FT	TRANSEM	564	582	S4 OF REPEAT II (POTENTIAL).
FT	DOMAIN	583	601	CYTOPLASMIC (POTENTIAL).
FT	TRANSEM	602	621	S5 OF REPEAT II (POTENTIAL).
FT	DOMAIN	622	674	EXTRACELLULAR (POTENTIAL).
FT	TRANSEM	675	699	S6 OF REPEAT II (POTENTIAL).
FT	DOMAIN	700	1148	CYTOPLASMIC (POTENTIAL).
FT	TRANSEM	1149	1166	S1 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1167	1182	EXTRACELLULAR (POTENTIAL).
FT	TRANSEM	1183	1202	S2 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1203	1214	CYTOPLASMIC (POTENTIAL).
FT	TRANSEM	1215	1233	S3 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1234	1243	EXTRACELLULAR (POTENTIAL).
FT	TRANSEM	1244	1262	S4 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1263	1281	CYTOPLASMIC (POTENTIAL).
FT	TRANSEM	1282	1301	S5 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1302	1388	EXTRACELLULAR (POTENTIAL).
FT	TRANSEM	1389	1413	S6 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1414	1468	CYTOPLASMIC (POTENTIAL).
FT	TRANSEM	1469	1487	S1 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1488	1502	EXTRACELLULAR (POTENTIAL).
FT	TRANSEM	1503	1522	S2 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1523	1530	CYTOPLASMIC (POTENTIAL).
FT	TRANSEM	1531	1549	S3 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1550	1558	EXTRACELLULAR (POTENTIAL).
FT	TRANSEM	1559	1577	S4 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1578	1596	CYTOPLASMIC (POTENTIAL).
FT	TRANSEM	1597	1616	S5 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1617	1680	EXTRACELLULAR (POTENTIAL).
FT	TRANSEM	1681	1705	S6 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1706	2326	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	1869	1873	POLY-GLN.
FT	DOMAIN	2040	2046	POLY-HIS.
FT	DOMAIN	371	388	BINDING TO THE BETA SUBUNIT (BY SIMILARITY).
FT	SITE	306	306	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).
FT	SITE	653	653	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).
FT	SITE	1362	1362	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).
FT	SITE	1650	1650	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).
FT	MOD_RES	1716	1716	PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT	CA_BIND	1734	1745	BY SIMILARITY.
FT	CARBOHYD	271	271	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1558	1558	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	406	406	D -> DDGLGIIVEPEQKPEDIQSVY (IN ISOFORM 2).
SQ	SEQUENCE	2326 AA;	264515 MW;	D58DEAA09E819B6B CRC64;

Query Match 5.0%; Score 92.5; DB 1; Length 2326;
 Best Local Similarity 23.2%; Pred. No. 20;
 Matches 80; Conservative 46; Mismatches 130; Indels 89; Gaps 17;

QY	60	NTILLPHIFFVPSHRKSIDMIISHLIHL-----LFTQAILVSLD	105
DB	1178	NVLKYLDYVETGVF-----TFEMVKINLGLILHPSYFRDLWNILDFIVSGALVAF	1232
QY	106	FFGSO---NTODRLYKVIVFLNKNVGRGLSICPLLVLQATISIFSIAKLKHP	161
DB	1333	FTSGRGKDLNTIKSLR-----VLRVLRPLK--TIKRLPKLKAIVDCVNSLKNVLNLI	1284
QY	162	SHILGFFLFSKV-LNMFIGVIFCCTLRPLPVK---RGQSSVCHTALFLFAHELHPQETVF	217
DB	1285	VYMLFMFIAFVIAVQLFKGKFFYCTDESKLEKDCRGOLVYDN-----DEIAEPREW	1338
QY	218	HTNDFEGCHLYRVHGPLKRLH-----GDYFTQITRGYLSAFTQACPRVSPVKRQAILL	273
DB	1339	KKCDF---HYDNVLWALITLFTVSTGEGWPTVLKNSIDATEDQGP--SPSYRMEISFY	1393

QY 274 LVSEV-----FTYWDFTSFSGVWTWINDSLVWLQVIVANSYAAI-----SPL 318
 Db 1394 VVVFVPEFFVNIIFVALIITFQEGDKVMSDCSL-----EKNERACIDFAISAKPL 1446
 QY 319 MLIYADNQIFKYLQ-MLFKFYLSPP-----KLMKF 348
 Db 1447 TRYPMQNK--QTFQYKMKFVSPFPFYLIMALIALNTIVLMKF 1489

RESULT 11
 GRPR_MOUSE STANDARD; PRT; 384 AA.
 AC P21729;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Gastrin-releasing peptide receptor (GRP-R) (GRP-preferring bombesin receptor).
 DE GRPR.
 GN Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-SWISS; TISSUE=Fibroblast;
 RX MEDLINE=91110536; PubMed=1671171;
 RA Battey J.F., Way J.M., Corjay M.H., Shapira H., Kusano K., Harkins R.,
 RA Wu J.M., Slattery T., Mann E., Feldman R.I.;
 RT "Molecular cloning of the bombesin/gastrin-releasing peptide receptor
 RT from Swiss 3T3 cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:395-399(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SWISS; TISSUE=Fibroblast;
 RX MEDLINE=91187004; PubMed=1707129;
 RA Giladi E.R., Giladi E., Brehm P., Goodman R.H., Segerson T.P.;
 RT "Cloning and functional characterization of a complementary DNA
 RT encoding the murine fibroblast bombesin/gastrin-releasing peptide
 RT receptor.";
 RL Mol. Endocrinol. 4:1956-1963(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93305484; PubMed=8391296;
 RA Giladi E., Nagalla S.R., Spindel E.R.;
 RT "Molecular cloning and characterization of receptors for the
 RT mammalian bombesin-like peptides.";
 RL J. Mol. Neurosci. 4:41-54(1993).
 CC -!- FUNCTION: RECEPTOR FOR GASTRIN-RELEASING PEPTIDE (GRP). THIS
 CC RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT
 CC ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: BRAIN (HYPOTHALAMUS), PANCREATIC ACINAR CELLS,
 CC AND FIBROBLASTS.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M57922; AAA75650.1; -;
 DR EMBL; M61000; AAA37744.1; -;
 DR PIR; A36553; A36553.
 DR PIR; A39003; A39003.
 DR GCRdb; GCR_0096; -;
 DR GCRdb; GCR_0097; -;
 DR MGD; MGI:95836; Grpr.
 DR InterPro; IPR000276; GPCR_Rhodopsin.

DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN_RECP_FL1; 1.
 DR PROSITE; PS0262; G-PROTEIN_RECP_FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 FT DOMAIN 1 39
 FT TRANSMEM 40 63
 FT DOMAIN 64 77
 FT TRANSMEM 78 97
 FT DOMAIN 98 115
 FT TRANSMEM 116 137
 FT DOMAIN 138 153
 FT TRANSMEM 154 175
 FT DOMAIN 176 209
 FT TRANSMEM 210 235
 FT DOMAIN 236 265
 FT TRANSMEM 266 286
 FT DOMAIN 287 299
 FT TRANSMEM 300 326
 FT DOMAIN 327 384
 FT CARBOHYD 5 5
 FT CARBOHYD 20 20
 FT CARBOHYD 24 24
 FT DISULFID 114 197
 FT LIPID 340 340
 FT CONFLICT 309 309
 SQ SEQUENCE 384 AA; 43214 MW; BF6D60387AA09A2C CRC64;

Query Match 5.0%; Score 92; DB 1; Length 384;
 Best Local Similarity 22.3%; Pred. NO. 3.6;
 Matches 60; Conservative 40; Mismatches 117; Indels 52; Gaps 10;

QY 40 PGFI-TIQIFVYPQASFGISANTILLPHIFTFVFSHRSKSDMIITSLHLLHLLFTQ 98
 Db 38 PGFIIVPAVYGLIIVIGLIGN--ITLIKIFCTVSMRNP-NLFISLLALGDLILLVTC 94
 QY 99 AILVSLDFGSGNTQDDLRKYVIVFLNKVMRGLSICITCLLSV--LQAIISPSIFSLAKL 156
 Db 95 APVDASKYLADRWLFRIGCKLIPFTQLTSGVSVFTLTALSADRYKAIVRPMDI----- 149
 QY 157 KPSASHILGFFLS-----WVLNMFIVGIFCCTLRLPP--VARGQSSVCHTALFLFAHEL 210
 Db 150 ----QASHALMKICLKAALIWIIVSMLLAIPAEVFSDLHPFHVDKTNQTFISCAPYHSNEL 206
 QY 211 HPQETVFTHTNDFEGCHLYRVHG-----PLKRLHGDYFIQITRGVLSAFTOP---- 256
 Db 207 HP-----KIHSMASFLVYVIPLAIIISVYFYFIARNLIQSAYNLPVEGN 250
 QY 257 --ACPRVSPVKRASQAILLVSVFVYVW 283
 Db 251 IHVKQIESKRKLAKTIVLVFVGLFAFCWL 279

RESULT 12
 FLOI_CRIGR STANDARD; PRT; 518 AA.
 AC P42557;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Folate transporter 1 (folate carrier protein) (Methotrexate uptake protein).
 DE protein.
 GN SLC19A1.
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetus.
 OX NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;

RX MEDLINE-94164933; PubMed-81119923;
RA Williams F.M.R., Murray R.C., Underhill T.M., Flintoff W.F.;
RT "Isolation of a hamster cDNA clone coding for a function involved in
RT methotrexate uptake";
RL J. Biol. Chem. 269:5810-5816(1994).
CC -!- FUNCTION: TRANSPORTER FOR THE INTAKE OF FOLATE. INVOLVED IN
CC METHOTREXATE UPTAKE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO THE SLC19A FAMILY OF TRANSPORTERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U03031; AAC52138.1; -;
DR InterPro: IPR002666; Folate_carrier.
DR Pfam: PF01770; Folate_carrier; 1.
KW Folate-binding; Transport; Transmembrane; Glycoprotein.
FT TRANSMEM 28 48
FT TRANSMEM 65 85
FT TRANSMEM 91 111
FT TRANSMEM 122 142
FT TRANSMEM 156 176
FT TRANSMEM 180 202
FT TRANSMEM 274 294
FT TRANSMEM 304 324
FT TRANSMEM 331 351
FT TRANSMEM 362 382
FT TRANSMEM 398 418
FT TRANSMEM 432 452
SQ SEQUENCE 518 AA; 58611 MW; 0B18267A134FC5AA CRC64;

Query Match 4.9%; Score 91; DB 1; Length 518;
Best Local Similarity 20.2%; Pred. No. 5.8;
Matches 76; Conservative 61; Mismatches 137; Indels 102; Gaps 17;

QY 28 LC-FNFPKMKILPGTITQIFYPQASFGISANT-----ILLFLHFTTFVFSH 75
DB 32 LCFEFGMAQLRPGESFITPYLL---QONFTIEQVTNEIIPVLSHLAVLPVIFLTDYL 88
QY 76 RKSIDMIHSLIHI---LLFTQAILVSL-DFFGQNTQDRLRYKVIIV-LNKVMRG 130
DB 89 RYKPI-LILOCLSFMCVWLLLLGTSVVMQLMVEVFSVTMAARIAYSSVIFSLVRPSRY 147
QY 131 LSCTPCLLSVLQAIISPSIFS--LAKLHPSASH-----ILGFPLFSWLNMFVGVIF 182
DB 148 QRMASYSRAAVLLGVFTSSVYLGQVLPLEQKSONSNMLNYISLGFITFSLGLSLFL---- 203
QY 183 CCTLRUPPKVR-----GSSVCHTALFLFAHELHP-----QETVFHT--NDFEG 224
DB 204 -----KRPKSLFFNKSALVHALPCELDQMPGPRPGKRLVSGCRNSFLV 254
QY 225 CHLYRVHGPKLRLHGD-----YFIQTIRGYLSAFTQACPRVSPVKRASQAILLVLSFVFT 280
DB 255 CMLSELVGNLRQPHVRLWCLWVFNAGY-----YLIYVYVHV 292
QY 281 YW-VDTTFPSGQVT-----WINDSLVWLVQVIVANSYAISPLMIYAD 324
DB 293 LWSIDKNLNYGAVDAASTLLSAITSFSGFKIRNALWSKLVASVIAQAGLVFCMYM 352
QY 325 NQIFKTLQMLWFKYLS 340
DB 353 VHYVTWVHKIWLVLMT 368

RESULT 13
YJ03_YEAST STANDARD; PRT; 555 AA.

AC P46996;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 61.5 kDa protein in TPK1-CCW7 intergenic region.
GN YJL163C OR J0344.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Obermaier B., Piravandi E., Rinke M., Domdey H.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z49438; CAA89458.1; -;
DR SGD: S0003699; YJL163C.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 84 104
FT TRANSMEM 126 146
FT TRANSMEM 189 209
FT TRANSMEM 230 250
FT TRANSMEM 258 278
FT TRANSMEM 357 377
FT TRANSMEM 387 407
FT TRANSMEM 429 449
FT TRANSMEM 460 480
FT TRANSMEM 492 512
FT TRANSMEM 524 544
SQ SEQUENCE 555 AA; 61524 MW; A72400A0B8B0A533 CRC64;

Query Match 4.9%; Score 91; DB 1; Length 555;
Best Local Similarity 22.0%; Pred. No. 6.2;
Matches 45; Conservative 34; Mismatches 64; Indels 62; Gaps 9;

QY 13 IMLFSLD-----LLFSTDILCFNFPKMKILPGTITQIFYPQASFGISA 59
DB 354 ILLVLDILFVCGTSCMPALILFST---YKWHAVELGYFISI-----LGIGR 400
QY 60 NTILL-----LFHIFTVFVSHRSKSID-----MIISHLSLHILLFTQ----- 98
DB 401 GVVLLVVSPTLLTKRIYOHNLNHSIDKIDFCIQFSMIVITLSLF-VMIRFGEKTPSM 459
QY 99 ---AIVLSDFGQSNTQDD-LRYKIVIVLNVKVRGLSICTPCLLSVLQAI----- 145
DB 460 IIPALQALSFAFCSPILQSGIIRYTKTKKGTGFMGAMALVRSCVMLVIPPILLKLGSTV 519
QY 146 -ISPSIFSLAKLKHPSASHILGFEL 169
DB 520 SVNPSLFMYIPFSTSIIVAILLTFFEL 544

RESULT 14
OIG1_HUMAN
ID OIG1_HUMAN STANDARD; PRT; 313 AA.
AC P47890; Q9UM76;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Olfactory receptor 1G1 (Olfactory receptor 17-209) (OR17-209).
GN OR1G1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20139433; PubMed=10673334;
 RA Glusman G., Sosinsky A., Ben-Asher E., Avidan N., Sonkin D., Bahar A.,
 RA Rosenthal A., Clifton S., Roe B., Ferraz C., Demalle J.G., Lancet D.,
 RT "Sequence, structure, and evolution of a complete human olfactory
 RT receptor gene cluster.";
 RL Genomics 63:227-245(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ferraz C., Demalle J.G.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 68-283 FROM N.A.
 RX MEDLINE=94272458; PubMed=8004088;
 RA Ben-Arie N., Lancet D., Taylor C., Khen M., Walker N.,
 RA Ledbetter D.H., Carrozzo R., Patel K., Sheer D., Lehrach H.,
 RA North M.A.;
 RT "Olfactory receptor gene cluster on human chromosome 17: possible
 RT duplication of an ancestral receptor repertoire.";
 RL Hum. Mol. Genet. 3:229-235(1994).
 CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF087928; AAF37317.1; -
 CC EMBL; U53583; AAC99556.1; -
 CC EMBL; U04689; AAA18352.1; -
 CC GCRDB; GCR_0856; -
 CC InterPro; IPR000276; GPCR_Rhodopsn.
 CC Pfam; PF00001; 7tm.1; 1.
 CC PRINTS; PR00237; GPCRHHODOPS
 CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
 CC PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Olfaction.
 FT DOMAIN 1 25 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 26 49 1 (POTENTIAL).
 FT DOMAIN 50 57 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 58 79 2 (POTENTIAL).
 FT DOMAIN 80 100 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 101 120 3 (POTENTIAL).
 FT DOMAIN 121 139 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 141 158 4 (POTENTIAL).
 FT DOMAIN 159 196 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 197 219 5 (POTENTIAL).
 FT DOMAIN 220 236 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 237 259 6 (POTENTIAL).
 FT DOMAIN 260 271 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 273 292 7 (POTENTIAL).
 FT DOMAIN 293 313 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 97 189 BY SIMILARITY.
 FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 95 95 S -> L (IN REF. 3).
 FT CONFLICT 125 125 A -> G (IN REF. 3).
 SQ SEQUENCE 313 AA; 34924 MW; 556C28EB731FD003 CRC64;
 Query Match 4.8%; Score 89.5; DB 1; Length 313;
 Best Local Similarity 23.3%; Pred. NO. 4.5;
 Matches 55; Conservative 32; Mismatches 74; Indels 75; Gaps 12;
 QY 10 MAEIMFLSLLLFSTD---ILCFNFPKMKILKCFITIQIFFFYPQASFGISANTILLF 66

Db 105 MLFWLEAFLLAVWAYDCYVAICHPHLHYILMSPL-----CIFLSVASWIMNAL 154
 QY 67 HIFTVFVSHRSKSIDMIISHLHILLHLLFTQAILVSLD---FFG-----SQNTQDDLRL 117
 Db 155 H-----SLHLLMNSLSFCANHEIPHFCDINPLLSLCTDPFT 194
 QY 118 YKVIVLKNKVRGSIICPCLLSVLQAIISPSIFSLAKLKHPSA-----SHILG 166
 Db 195 NELVIFITGGLTGL-ICVLCIL-----ISYTNVFSTI-LKIPSAQGRKKAFTSCSHLS- 246
 QY 167 FFLFSWLNMEIGVIFCCTLRLPPVKRCQ-----SSVCHTA-----LFLFAHELHPQE 214
 Db 247 -----VVSFLFGTSFCVDFSPSTHSAQKDTVASVMYTVVTPMLNPFYSLRNOE 296
 RESULT 15
 GRPR_HUMAN STANDARD; PRT; 384 AA.
 ID GRPR_HUMAN AC P30550;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gastrin-releasing peptide receptor (GRP-R) (GRP-preferring bombesin
 DE receptor).
 DE GRPR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=92011639; PubMed=1655761;
 RA Corjay M.H., Dobrzanski D.J., Way J.M., Viallet J., Shapira H.,
 RA Worland P., Sausville E.A., Battley J.F.;
 RT "Two distinct bombesin receptor subtypes are expressed and functional
 RT in human lung carcinoma cells.";
 RL J. Biol. Chem. 266:18771-18779(1991).
 CC -1- FUNCTION: RECEPTOR FOR GASTRIN BY ASSOCIATION WITH G PROTEINS THAT
 CC RECEPTOR MEDIATES ITS ACTION BY ACTIVATING CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M73481; AAA8050.1; -
 CC PIR; A41007; A41007.
 CC GCRDB; GCR_0187; -
 CC MIM; 305670; -
 CC InterPro; IPR000276; GPCR_Rhodopsn.
 CC Pfam; PF00001; 7tm.1; 1.
 CC PRINTS; PR00237; GPCRHHODOPS
 CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 CC PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Lipoprotein; Palmitate.
 FT DOMAIN 1 38 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 39 62 1 (POTENTIAL).
 FT DOMAIN 63 76 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 77 96 2 (POTENTIAL).
 FT DOMAIN 97 114 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 115 136 3 (POTENTIAL).
 FT DOMAIN 137 152 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 153 174 4 (POTENTIAL).
 FT DOMAIN 175 208 EXTRACELLULAR (POTENTIAL).

	TRANSHEM	209	234	5 (POTENTIAL).	
FT	DOMAIN	235	264	CYTOPLASMIC (POTENTIAL).	
FT	TRANSHEM	265	285	6 (POTENTIAL).	
FT	DOMAIN	286	298	EXTRACELLULAR (POTENTIAL).	
FT	TRANSHEM	299	325	7 (POTENTIAL).	
FT	DOMAIN	326	384	CYTOPLASMIC (POTENTIAL).	
FT	CARBOHYD	20	20	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	DISULFID	113	196	BY SIMILARITY.	
FT	LIPID	339	339	PALMITATE (BY SIMILARITY).	
SQ	SEQUENCE	384 AA;	43198 MW;	70A75FLI1D9C82FI9 CRC64;	
	Query Match	4.8%;	Score 89.5;	DB 1; Length 384;	
	Best Local Similarity	21.3%;	Pred. No. 5.6;		
	Matches 55;	Conservative	37;	Mismatches 103; Indels 63; Gaps 10;	
Qy	56 GISANTILLPHIFTFVFSHRKSIDMIISHLSIHILLLETTQAIIYLVLDFGSGNTODD	115	:	:	:
Db	54 GLIGN--ITLIKIFICTVKSMRNP-NLFISSIALGDLLLTTCAPVDASRYLADRWLFGR	110	:	:	:
Qy	116 LRVKIVVFNLKVMRGSLSCTCLLSV--LOAIISPSTFSLAKLHPASHTLGFF----	169	:	:	:
Db	111 IGCKLIPIQLTSGVSFTLTALSADRYKAIVRPMDI-----QASHALKICKAA	162	:	:	:
Qy	170 FSWLNFMFIGVFCTLRLPVPKRQSVCHTALFLFAHELHPQTVEHPTWDFEGCHLY-	228	:	:	:
Db	163 FIWIISMLLAI-----PEAVESDLHPFHSEST---NQTEISCAPYP	200	:	:	:
Qy	229 -----RVHG-----PKRLKHGXFIQTIIRGVLSAFTOP-----ACPRVSPVK	265	:	:	:
Db	201 HSNELHPKIHSMASFLVFYVIPLSIIISVVYIFIANKLIOQSAYNLPVEGNHVHKQIESRK	260	:	:	:
Qy	266 RASAQILLLVSFVTYYW	283	:	:	:
Db	261 RLAKTLVLFVGFLAFACWL	278	:	:	:

Search completed: September 18, 2002, 10:56:40
Job time: 246 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 18, 2002, 10:52:14 ; Search time 28.4 Seconds
(without alignments)
2174.619 Million cell updates/sec

Title: US-09-728-309-3
Perfect score: 1855
Sequence: 1 MLKLVIIENAEIMFLSLD.....YLSPPKMLKFNROCGSTKK 357

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	406.5	21.9	299	11	Q9EQ40 mus musculus
2	388.5	20.9	299	11	Q9EQ37 mus musculus
3	386.5	20.8	300	11	Q9EQ38 mus musculus
4	365.5	19.7	310	11	Q62855 rattus norv
5	361	19.5	312	11	Q9EQ42 mus musculus
6	358.5	19.3	303	11	Q9EQ41 mus musculus
7	349.5	18.8	302	11	Q9EQ35 mus musculus
8	347.5	18.7	303	11	Q9EQ36 mus musculus
9	334.5	18.0	303	11	Q9EPA4 mus musculus
10	334.5	18.0	305	11	Q92195 mus musculus
11	330	17.8	313	11	Q9EQ49 mus musculus
12	329.5	17.8	311	11	Q62852 rattus norv
13	322.5	17.4	310	11	Q9EQ46 mus musculus
14	320.5	17.3	309	11	Q9EP79 mus musculus
15	320.5	17.3	310	11	Q9EQ44 mus musculus
16	313.5	16.9	315	11	Q62850 rattus norv

17	312.5	16.8	321	11	Q62856	Q62856 rattus norv
18	309	16.7	309	11	Q9EQ45	Q9EQ45 mus musculus
19	306.5	16.5	310	11	Q9EP93	Q9EP93 mus musculus
20	305.5	16.5	278	11	Q9EQ39	Q9EQ39 mus musculus
21	305	16.4	331	11	Q9EQ43	Q9EQ43 mus musculus
22	303.5	16.4	310	11	Q9EQ47	Q9EQ47 mus musculus
23	301.5	16.3	276	11	Q9EQ34	Q9EQ34 mus musculus
24	298.5	16.1	303	11	Q92196	Q92196 mus musculus
25	291.5	15.7	308	11	Q9WU03	Q9WU03 mus musculus
26	291	15.7	310	11	Q9EP51	Q9EP51 mus musculus
27	289	15.6	313	11	Q9EQ50	Q9EQ50 mus musculus
28	284.5	15.3	310	11	Q9EQ51	Q9EQ51 mus musculus
29	279.5	15.1	302	11	Q9EQ52	Q9EQ52 mus musculus
30	275.5	14.9	273	11	Q62854	Q62854 rattus norv
31	274	14.8	279	11	Q9EQ48	Q9EQ48 mus musculus
32	269.5	14.5	335	11	Q62853	Q62853 rattus norv
33	268.5	14.5	273	11	Q62851	Q62851 rattus norv
34	268.5	14.5	310	11	Q9WU02	Q9WU02 mus musculus
35	265.5	14.3	310	11	Q9WU01	Q9WU01 mus musculus
36	265.5	14.3	315	11	Q9EPB8	Q9EPB8 mus musculus
37	238	12.6	349	4	Q9BX81	Q9BX81 homo sapien
38	236	12.7	334	4	Q9BXE8	Q9BXE8 homo sapien
39	192.5	10.4	353	4	Q9G2P7	Q9G2P7 homo sapien
40	184	9.9	311	4	Q9BXE9	Q9BXE9 homo sapien
41	183.5	9.9	313	11	Q9EPS4	Q9EPS4 mus musculus
42	175.5	9.5	332	11	Q9EPS7	Q9EPS7 mus musculus
43	166.5	9.0	300	11	Q9EPT1	Q9EPT1 mus musculus
44	164.5	8.9	300	11	Q9EPS8	Q9EPS8 mus musculus
45	162.5	8.6	298	11	Q9EPT0	Q9EPT0 mus musculus

ALIGNMENTS

RESULT 1

ID	Q9EQ40	PRELIMINARY;	PRT;	299 AA.
AC	Q9EQ40;			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	VOMERONASAL RECEPTOR VIRC3.			
GN	VIRC3.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/SVJ;			
RX	MEDLINE=20568485; PubMed=11116090;			
RA	Del Punta K., Rothman A., Rodriguez I., Mombaerts P.;			
RT	"Sequence diversity and genomic organization of vomeronasal receptor			
RT	genes in the mouse.";			
RL	Genome Res. 10:1958-1967(2000).			
DR	EMBL; AF291499; AAG42093.1; -.			
DR	MGI; MGI:2148524; Virc3.			
DR	InterPro; IPR000276; GPCR_Rhodpsn.			
DR	PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.			
KW	Receptor.			
SQ	SEQUENCE 299 AA; 34603 MW; 9693907794F3D105 CRC64;			

Query Match 21.9%; Score 406.5; DB 11; Length 299;

Best Local Similarity 34.2%; Pred. No. 3.8e-29;

Matches 107; Conservative 47; Mismatches 110; Indels 49; Gaps 9;

Oy 50 YPOASFGISANTILLFIHFTEVFSHRKSIDMIITSLHILLFTQAILVLDFFGS 109

Db 9 YIQAGLVIANCLLVFYIF-MVLGHRPKPMDLISCOQTFIHMLFFTAGDILHTDIES 67

Oy 110 QNTQDRLRYKIVFVFNKVMYRGLSCTPCLLSVLQAI-TSPSIFSLAKLHKPSASHILGFF 168

```
Db 68 MNIENDFCKTTFYICRWVRGLSICCTCLLSVFQAVTISPNTSLAKFKHLKKYTIINAF 127
Qy 169 LFSWVLNM-----FIGVI-----FCCTLRLLPPVKRGSSVCHTA--LFL 205
Db 128 FYIWSFNLSFSSNLFIYVGAYTNVSETNQMKYKCSLFPNYYIIRGLLITVTSRDVFL 187
Qy 206 FAHELHPQETVHTWDFEGCHLYRVHGPKLRHLGDFYQTIOTINGYLSAFTQACPRVSPVK 265
Db 188 VG-----VMLITSTYVVIILFRHQCKHLH-----SIRHL-----RASPEK 224
Qy 266 RASQAILLV-SFVFTYVWDFTFSGGVTWINDSLLVQVIVANSYAAISPLMIYAD 324
Db 225 RATQILLVLIYFVVVYVWDFIISSTSVLLMWDPVILTQKFMNAYPTITPLVQISSD 284
Qy 325 NQIKFTLQMLWFK 337
Db 285 NRIINLKNLQSK 297

RESULT 2
Qy 99037 PRELIMINARY; PRT; 299 AA.
AC Q99037;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VOMERONASAL RECEPTOR V1RC6.
GN V1RC6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20568485; PubMed=11116090;
RA Del Punta K., Rothman A., Rodriguez I., Mombaerts P.;
RT "Sequence diversity and genomic organization of vomeronasal receptor
RT genes in the mouse.";
RL Genome Res. 10:1958-1967(2000).
DR EMBL; AF291502; AAG42096.1; -.
DR MGD; MGI:2148527; V1rc6.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR PROSITE; PS50262; G_PROTEIN_RECPE_FL_2; 1.
KW Receptor.
SQ SEQUENCE 299 AA; 34423 MW; 86CF4F75937B7A27 CRC64;

Query Match 20.9%; Score 388.5; DB 11; Length 299;
Best Local Similarity 31.9%; Pred. No. 1.7e-27;
Matches 101; Conservative 51; Mismatches 120; Indels 45; Gaps 7;

Qy 44 TIQIFYPOASFGISANTILLFHIFTVFVSHRSKSIDMIISHLIHLILFTQAILVS 103
Db 3 SVENLYFQAGFGLANFLLYFYIF-ILGHRPKMDLISCOLTLVHILMELTGNLWL 61
Qy 104 LDFFGSQNTDRLRYKIVFLNKNVRGLSICTPCLLSVLOAI-ISPSTFSLAKLHPSAS 162
Db 62 ADIFGLAENDIKCKATFYTSRVVRGLSICITCLLSVFOAVTISPSTSLAKPKQLKK 121
Qy 163 HILGFFLFSWVLNMFIVGIFCCTLRLLPPVKRGSSVCHTAFLFAHELHPQETVHTWDF 222
Db 122 YNNALFYIMFNL-----SVCNNLFFVGGFTNVSETK-QVKVT 160
Qy 223 ECHLYRVHGPKLR-----HGDFYFIQTIRGYSALT-----QPACPRV 261
Db 161 KSCSLFPNYYIIRLIVLSYSDVFLVAVMLTSSAYVNNILHRQRYKHLHSKSLRT 220
Qy 262 SPVKRASQAILL-VSFVFTYVWDFTFSGGVTWINDSLLVQVIVANSYAAISPLML 320
Db 221 SPEKKATQITILLVVFFVVMYVWDFIISFTSLLMMDPVVLTQKFLIYAYPTITPLVQ 280
Qy 321 IYADNQIKFTLQMLWFK 337
```

```
Db 281 ISSDKRIINVLKNLQSK 297

RESULT 3
Qy 99038 PRELIMINARY; PRT; 300 AA.
AC Q99038;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VOMERONASAL RECEPTOR V1RC5.
GN V1RC5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20568485; PubMed=11116090;
RA Del Punta K., Rothman A., Rodriguez I., Mombaerts P.;
RT "Sequence diversity and genomic organization of vomeronasal receptor
RT genes in the mouse.";
RL Genome Res. 10:1958-1967(2000).
DR EMBL; AF291501; AAG42095.1; -.
DR MGD; MGI:2148526; V1rc5.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR PROSITE; PS50262; G_PROTEIN_RECPE_FL_2; 1.
KW Receptor.
SQ SEQUENCE 300 AA; 34330 MW; 2D69A8680ABCD433 CRC64;

Query Match 20.8%; Score 386.5; DB 11; Length 300;
Best Local Similarity 32.7%; Pred. No. 2.5e-27;
Matches 104; Conservative 48; Mismatches 113; Indels 53; Gaps 9;

Qy 44 TIQIFYPOASFGISANTILLFHIFTVFVSHRSKSIDMIISHLIHLILFTQAILVS 103
Db 3 SLENVLYFQAGFGLANFLLYFYIF-ILGHRPKLDLISCOLTFVHILMILTGNNVLM 61
Qy 104 LDFFGSQNTDRLRYKIVFLNKNVRGLSICTPCLLSVLOAI-ISPSTFSLAKLHPSAS 162
Db 62 SDIFESLNVENDIKCKATLYTNVRMGLSISITCLLSVIAQVITISPSTMLAKFKHLKK 121
Qy 163 HIL-GFFLFSWVLNMFIVGIF-----CCTLRLLPPVKRGSSVCH 200
Db 122 HMVNASFYIWSFNLSLILFYTGFTNVSETKQMKITKCSILPMNYIIRGHVVT 181
Qy 201 TA--LFLFAHELHPQETVHTWDFEGCHLYRVHGPKLRHLGDFYFIQTIRGYSALTQAC 258
Db 182 TVRDVFLVG-----VMLITSAYNVIILFRHEROCKHLHSISHL----- 219
Qy 259 PRVSKRASQAILL-VSFVFTYVWDFTFSGGVTWINDSLLVQVIVANSYAAISPL 317
Db 220 -RASEKKATQITILLVVFFVVMYVWDFILSSTSVTLWMDPVILTQKFLMNAIPIITP 278
Qy 318 LMLIYADNQ---IFEKTLQ 332
Db 279 LLQISSDKRINVMKTLQ 296

RESULT 4
Qy 99038 PRELIMINARY; PRT; 310 AA.
AC Q99038;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PHEROMONE RECEPTOR VN6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
```

```
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=VOMERONASAL ORGAN;
RX MEDLINE=96028094; PubMed=7585937;
RA Dulac C., Axel R.;
RT "A novel family of genes encoding putative pheromone receptors in
mammals.";
RL Cell 83:195-206(1995).
DR EMBL: U36898; AAC52287.1;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR004072; Vomeron1_receptor.
DR PRINTS: PR01534; VOMERONASLIR.
DR PROSITE: PS0262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
SQ SEQUENCE 310 AA; 35301 MW; 50CB3BA274A03CFD CRC64;

Query Match 19.7%; Score 365.5; DB 11; Length 310;
Best Local Similarity 32.1%; Pred. No. 2.1e-25;
Matches 104; Conservative 54; Mismatches 111; Indels 55; Gaps 10;

QY 39 LPGFITIQIPYPOASGISANTILLFLPHFTFVFSHRKSIDMIISHLSLIHILLFTQ 98
DB 7 LYGVVDQQAIFSFVWIGISFNLSLFHFIFQFLERLRITDIIISLLAHILGLMTVM 66
QY 99 AILVSLDFQSQNTQDRLRYKVIVFLNKMVRLGSLCTPCLLSVLQAI-IPSFISFLAKLK 157
DB 67 GFR-AVDIFASQVWMDIKSLAHLRLRLGLSLCATCLLSIFQAITLSRSSCLAKFK 125
QY 158 HPSASHILGFLFSWLNMFVIFCTCLRLPPVKRGQSSVCHTALFLFAHELHPQETVP 217
DB 126 YKSTQHSCLSLVLWAFYMSGTHYSFTI-----VADYNFSSRLIF 167
QY 218 HTNDFEGCHLYRVHGPLKRLHGD-YFIOTI-----RGYLSAF----- 253
DB 168 VT-----ESCIIIL----PMDYITRDLFFILGIFPROVSFGLMALSSGYVALLCRHKQAQH 220
QY 254 --TOPACPRVPVKRASQAIIILLYS-FVFTYVWDFTFSGGVTWINDSLLVWLQVIVAN 310
DB 221 LHRTSLSPKASPEQARTITILLMSFFVLMVCLDCTISAS-RLMHNGEPIHHSIQMVSN 279
QY 311 SYAISPMLIYADNQIFKTLQML 334
DB 280 SYATLSPLLIIVTENRISRLKSL 303

RESULT 5
Q9EQ42 PRELIMINARY; PRT; 312 AA.
ID Q9EQ42
AC Q9EQ42
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VOMERONASAL RECEPTOR VIRCI.
GN VIRCI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20568485; PubMed=11116090;
RA Del Punta K., Rothman A., Rodriguez I., Mombaerts P.;
RT "Sequence diversity and genomic organization of vomeronasal receptor
genes in the mouse.";
RL Genome Res. 10:1958-1967(2000).
DR EMBL: AF291497; AAG42091.1;
DR MGI:2148522; Vircl.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR PROSITE: PS0262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.

QY 50 YPQASFGISANTILLFLPHFTFVFSHRKSIDMIISHLSLIHILLFTQAILVSLDFEGS 109
DB 9 YFQAGLGLVANMVLLEFYIF-MSWVTRPKPTDILSCQLTFVPHNGLSTGGDILLTDLFE 67
QY 109 SONTODDLRYKVIVFLNKMVRLGSLCTPCLLSVLQAI-IPSFISFLAKLKHPSSHILGF 167
DB 68 LLNIENDLKCKTIFYISRVMRGLSICCTCLLSVFQAVTISPTSLAKFKQKLKMYVCV 127
QY 168 FLFSWLNMFVGV--IF-----CCTLRLLPPVKRGQSSVCHTA--LF 204
DB 128 FLCIWSFNLAFTNRIFYVGGFTNVSETNQMVTKSCSLPMNIIIRGLITISTSRDVF 187
QY 205 LFAHELHPQETVHTNDFEGCHLYRVHGPLKRLHGDYFIQIRGYLSAFTQACPRVSPV 264
DB 188 LVG-----VMLTTSVVMVITMFRYHROCKYL---YSISHL-----RESPE 224
QY 265 KRASQAIIILLYS-FVFTYVWDFTFSGGVTWINDSLLVWLQVIVANSYAISPMLIYA 323
DB 225 KRATQTILLVSEFVVMYVWDFIISFTSDMIWYDPLILTQVKFMVAYAITPLVQISS 284
QY 324 DNOIFKTLQMLWFKY 338
DB 285 DNRILIMLNKLNOSKH 299

RESULT 6
Q9EQ41 PRELIMINARY; PRT; 303 AA.
ID Q9EQ41
AC Q9EQ41
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VOMERONASAL RECEPTOR VIRC2.
GN VIRC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20568485; PubMed=11116090;
RA Del Punta K., Rothman A., Rodriguez I., Mombaerts P.;
RT "Sequence diversity and genomic organization of vomeronasal receptor
genes in the mouse.";
RL Genome Res. 10:1958-1967(2000).
DR EMBL: AF291498; AAG42092.1;
DR MGI:2148523; Virc2.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR PROSITE: PS0262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
SQ SEQUENCE 303 AA; 35061 MW; BA6134B99EEFOAB0 CRC64;

Query Match 19.3%; Score 358.5; DB 11; Length 303;
Best Local Similarity 31.8%; Pred. No. 9e-25;
Matches 100; Conservative 51; Mismatches 114; Indels 49; Gaps 9;

QY 50 YPQASFGISANTILLFLPHFTFVFSHRKSIDMIISHLSLIHILLFTQAILVSLDFEGS 109
DB 9 YFQAGLGLVANMVLLEFYIF-IILGHRPKPTDILSCQLTFVHIMMFLAGGDITWIDTIFET 67
QY 110 QNTODDLRYKVIVFLNKMVRLGSLCTPCLLSVLQAI-IPSFISFLAKLKHPSSHILGF 168
DB 68 LNIENTDFCKTIFYISRVMRGLSICNTCLLSVQAVTISPTSLAKFKPKLKYIYAF 127
```

```
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=VOMERONASAL ORGAN;
RX MEDLINE=96028094; PubMed=7585937;
RA Dulac C., Axel R.;
RT "A novel family of genes encoding putative pheromone receptors in
mammals.";
RL Cell 83:195-206(1995).
DR EMBL: U36898; AAC52287.1;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR004072; Vomeron1_receptor.
DR PRINTS: PR01534; VOMERONASLIR.
DR PROSITE: PS0262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
SQ SEQUENCE 310 AA; 35301 MW; 50CB3BA274A03CFD CRC64;

Query Match 19.5%; Score 361; DB 11; Length 312;
Best Local Similarity 34.0%; Pred. No. 5.5e-25;
Matches 107; Conservative 47; Mismatches 111; Indels 50; Gaps 10;

QY 50 YPQASFGISANTILLFLPHFTFVFSHRKSIDMIISHLSLIHILLFTQAILVSLDFEGS 108
DB 9 YFQAGLGLVANMVLLEFYIF-MSWVTRPKPTDILSCQLTFVPHNGLSTGGDILLTDLFE 67
QY 109 SONTODDLRYKVIVFLNKMVRLGSLCTPCLLSVLQAI-IPSFISFLAKLKHPSSHILGF 167
DB 68 LLNIENDLKCKTIFYISRVMRGLSICCTCLLSVFQAVTISPTSLAKFKQKLKMYVCV 127
QY 168 FLFSWLNMFVGV--IF-----CCTLRLLPPVKRGQSSVCHTA--LF 204
DB 128 FLCIWSFNLAFTNRIFYVGGFTNVSETNQMVTKSCSLPMNIIIRGLITISTSRDVF 187
QY 205 LFAHELHPQETVHTNDFEGCHLYRVHGPLKRLHGDYFIQIRGYLSAFTQACPRVSPV 264
DB 188 LVG-----VMLTTSVVMVITMFRYHROCKYL---YSISHL-----RESPE 224
QY 265 KRASQAIIILLYS-FVFTYVWDFTFSGGVTWINDSLLVWLQVIVANSYAISPMLIYA 323
DB 225 KRATQTILLVSEFVVMYVWDFIISFTSDMIWYDPLILTQVKFMVAYAITPLVQISS 284
QY 324 DNOIFKTLQMLWFKY 338
DB 285 DNRILIMLNKLNOSKH 299

RESULT 6
Q9EQ41 PRELIMINARY; PRT; 303 AA.
ID Q9EQ41
AC Q9EQ41
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VOMERONASAL RECEPTOR VIRC2.
GN VIRC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20568485; PubMed=11116090;
RA Del Punta K., Rothman A., Rodriguez I., Mombaerts P.;
RT "Sequence diversity and genomic organization of vomeronasal receptor
genes in the mouse.";
RL Genome Res. 10:1958-1967(2000).
DR EMBL: AF291498; AAG42092.1;
DR MGI:2148523; Virc2.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR PROSITE: PS0262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
SQ SEQUENCE 303 AA; 35061 MW; BA6134B99EEFOAB0 CRC64;

Query Match 19.3%; Score 358.5; DB 11; Length 303;
Best Local Similarity 31.8%; Pred. No. 9e-25;
Matches 100; Conservative 51; Mismatches 114; Indels 49; Gaps 9;

QY 50 YPQASFGISANTILLFLPHFTFVFSHRKSIDMIISHLSLIHILLFTQAILVSLDFEGS 109
DB 9 YFQAGLGLVANMVLLEFYIF-IILGHRPKPTDILSCQLTFVHIMMFLAGGDITWIDTIFET 67
QY 110 QNTODDLRYKVIVFLNKMVRLGSLCTPCLLSVLQAI-IPSFISFLAKLKHPSSHILGF 168
DB 68 LNIENTDFCKTIFYISRVMRGLSICNTCLLSVQAVTISPTSLAKFKPKLKYIYAF 127
```

```
QY 169 LFSWLN-----MFIG-----VIFCCTLRPPVKRGQSSVCHTALFL---- 205
Db 128 FCIWSENLSSNQILYVGGFTNLSETNQKVTKCTL-LP-----KNYIIKGLILTVS 180
QY 206 FAHELHPQETVHTNDFECGCHLYRVHGPPLKRLHGDYFIQTIIRGYLSAFTQACPRVSPVK 265
Db 181 ISRDVFLVGVMILTSTYMLNILFRHQROCKHLHSYSL-----RASPEK 224
QY 266 RASQAAILLL-VSFVFTYVWDFTFSGGGVTWINDSLVWLQVIVANSYAAISPLMLIVAD 324
Db 225 RATQTTILLVVFVVMYVWDFIISSTSVMLWMDPVLTVQRFVCMAYPTIIPLVQIISD 284
QY 325 NOIFKTLQMLWFKY 338
Db 285 NRIIMLENNQSKH 298

RESULT 7
QYEQ35 PRELIMINARY; PRT; 302 AA.
AC Q9EQ35;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VOMERONASAL RECEPTOR VIRC8.
GN VIRC8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20568485; PubMed=11116090;
RA Del Punta K., Rothman A., Rodriguez I., Mombaerts P.;
RT "Sequence diversity and genomic organization of vomeronasal receptor
RT genes in the mouse.";
RL Genome Res. 10:1958-1967(2000).
DR EMBL; AF291504; AAG42098.1; -.
DR MGD; MGI:2148529; VirC8.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR PROSITE; PS50262; G_PROTEIN_RECEPTOR_F1_2; 1.
KW Receptor.
SQ SEQUENCE 302 AA; 34790 MW; D8435297F49750D4 CRC64;

Query Match 18.8%; Score 349.5; DB 11; Length 302;
Best Local Similarity 34.6%; Pred. No. 5,9e-24;
Matches 110; Conservative 40; Mismatches 109; Indels 59; Gaps 12;

QY 50 YPOASFGISANTILLFHIFTVFVSHRSKSIDMISHLSLHILL-----FTQAILVS 103
Db 9 YFOAGLGILANTFLLCFYFT-ILCHRSKPMDLTSCQLTLAHLFLLAGDNNLAVILES 67
QY 104 LDFFGQNTQDDLRYKIVVFLNKMVRLGSLCTICPCLLSVLQAI-ISPISFLAKLKHPSAS 162
Db 68 L-----NIENDPKKAIYTRNRMVGRGSGICITCLLSVFOAVTISPTSLAKFKHKLK 121
QY 163 HILGFLFSWLNMFIVFCCTLRPPVKRGQSSVCHT-----ALFLFAH---EL 210
Db 122 YIIYAVFYIWFNFLV-----CSHLLVYV-GFNNVSETNQMRVTESCSFLMNYIISGL 175
QY 211 HPQETVFH-----TNDFEGCHLYRVHGPPLKRLHGDYFIQTIIRGYLSAFTQACPR 260
Db 176 ILTVTTFRDVLVGVMLITNAVYIILFRHQHCKHLH-----SIRHL-----R 219
QY 261 VSPVKRASQAAILLL-VSFVFTYVWDFTFSGGGVTWINDSLVWLQVIVANSYAAISPLM 319
Db 220 VSPEKRATQITLLMVFVVMYVWDFIISSTSVMLWMDPVLTVQRFVCMAYPTIIPLV 279
QY 320 LIYADNOIFKTLQMLWFK 337
Db 280 QISSDKRVINVLKNLQSK 297
```

```
RESULT 8
QYEQ36 PRELIMINARY; PRT; 303 AA.
AC Q9EQ36;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VOMERONASAL RECEPTOR VIRC7.
GN VIRC7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20568485; PubMed=11116090;
RA Del Punta K., Rothman A., Rodriguez I., Mombaerts P.;
RT "Sequence diversity and genomic organization of vomeronasal receptor
RT genes in the mouse.";
RL Genome Res. 10:1958-1967(2000).
DR EMBL; AF291503; AAG42097.1; -.
DR MGD; MGI:2148528; VirC7.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR PROSITE; PS50262; G_PROTEIN_RECEPTOR_F1_2; 1.
KW Receptor.
SQ SEQUENCE 303 AA; 34965 MW; 8B2C9EE34C71DDE7 CRC64;

Query Match 18.7%; Score 347.5; DB 11; Length 303;
Best Local Similarity 28.9%; Pred. No. 9e-24;
Matches 90; Conservative 56; Mismatches 120; Indels 45; Gaps 7;

QY 45 IQIFYPQASFGISANTILLFHIFTVFVSHRSKSIDMISHLSLHILLFTQAILVSL 104
Db 11 LENILYFOAGLGVLANIFLLFYIF-IIIGHRPKMDLISCOLTFVHNMFLAGENFWLA 69
QY 105 DFFGQNTQDDLRYKIVVFLNKMVRLGSLCTICPCLLSVLQAI-ISPISFLAKLKHPSASH 163
Db 70 NIFESLNVENDEFKCKATFYTKRMVRLGSLCITCLLSVFOAVMISHRNSLLAKFKHLKIY 129
QY 164 ILGFLFSWLNMF1--GVIF-----CCTLRPPVKRGQSSVCHTA 202
Db 130 MINALFYIWTENLSLSSNLIFYVGGFTNMSQTKQLKVTKSCSLIPMNYIIRGL-----IL 184
QY 203 LFLFAHELHPQETVHTNDFEGCHLYRVHGPPLKRLHGDYFIQTIIRGYLSAFTQACPRVS 262
Db 185 IVISSRDMFLVGVMLTSTYMWIILCRHQROCKHLHSTSHL-----RAS 228
QY 263 PVKRASQAAILLL-VSFVFTYVWDFTFSGGGVTWINDSLVWLQVIVANSYAAISPLMLI 321
Db 229 PEKRATQITLLVVFVVMYVWDFIISSTSVMLWMDPVLMIVQKLVVNAVYPAITPFVQI 288
QY 322 YADNOIFKTLQ 332
Db 289 SSDTRVIRVVK 299

RESULT 9
QYEQ36 PRELIMINARY; PRT; 303 AA.
AC Q9EQ36;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VN12 (VOMERONASAL RECEPTOR VIRAL).
GN VIRAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```



```

RN SEQUENCE FROM N.A.
RP Lane R.P., Rowen L., Friedman C., Trask B.J., Hood L.;
RT "Genomic characterization of the murine chromosome 6D1 VNO olfactory
RT receptor cluster";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=129/SVJ;
RX MEDLINE=20568485; PubMed=11116090;
RA Del Punta K., Rothman A., Rodriguez I., Mombaerts P.;
RT "Sequence Diversity and Genomic Organization of Vomeronasal Receptor
RT Genes in the Mouse.";
RL Genome Res. 10:1958-1967(2000).
DR EMBL; AF129005; AAC43249.1; -.
DR EMBL; AF291481; AAC42075.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 303 AA; 34612 MW; 4FC872D15F9AD6C8 CRC64;

Query Match 18.0%; Score 334.5; DB 11; Length 303;
Best Local Similarity 31.2%; Pred. No. 1.4e-22;
Matches 97; Conservative 46; Mismatches 115; Indels 53; Gaps 8;

QY 49 FYQASFGISANTILLFHIFTVFHSRKSIDMIISHLSLHILLFTQAILVSLDFFG 108
DB 17 FFSEIGIGISGNSFLFLHILKFIHGHRSLSDLPICGLLSLHLLMLLVMA-FIATDIFI 75
QY 109 SONTQDDLRKYVFLNKMVGRGLSICPCLLSVLQAI-SPSIFSLAKLKHPSASHILGF 167
DB 76 SWRGWDDIICKFLVLYRVLRGLSLCTTSMLSVLQAILSPRSCLAKFKRKSLLHISCA 135
QY 168 FLFSWLNMFIVGFCCTLRPPVKRGQSSVCHTALFLFAHELHPQETVFTNDF---E 223
DB 136 ILFSLVLYMLIGSQLSVIIATP-----DYFIOTI-----RGYLSAF-----TO-----PACPRV 170
QY 224 GCHLYRVHGPLKRLHG-----DYFIOTI-----RGYLSAF-----TO-----PACPRV 261
DB 171 SCSTLPUSYVQSMFSTLLVIRDVFLISLWVSTWYVALLCRRHKKTQHLQGISLSPKT 230
QY 262 SPVKRASQAILLVSVFVFTVWDFTFSGGVTVWINDSLVWLQVIVANSYAAISPLMLI 321
DB 231 SPKQATQTLMLMSFFVLMTIYDVIVSCSRTMFLNDPTSYNMQIFVYVHIYATVSPFVM 290
QY 322 YADNQIEKTLQ 332
DB 291 STEKHIVNCLR 301

RESULT 10
Q92195 Q92195 PRELIMINARY; PRT; 305 AA.
AC Q92195;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PHEROMONE RECEPTOR 1.
GN VIRAL OR MPRI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C; TISSUE=VOMERONASAL RECEPTOR;
RA Saito H.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12725; CAA73257.1; -.
DR MGD; MGI:1333759; V1ral.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR004072; Vomeron1_receptor.

```

```

DR PRINTS; PRO1534; VOMERONASLIR.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor; G-protein coupled receptor.
SQ SEQUENCE 305 AA; 34874 MW; C0B5F0872DCC8F4D CRC64;

Query Match 18.0%; Score 334.5; DB 11; Length 305;
Best Local Similarity 31.2%; Pred. No. 1.4e-22;
Matches 97; Conservative 46; Mismatches 115; Indels 53; Gaps 8;

QY 49 FYQASFGISANTILLFHIFTVFHSRKSIDMIISHLSLHILLFTQAILVSLDFFG 108
DB 19 FFSEIGIGISGNSFLFLHILKFIHGHRSLSDLPICGLLSLHLLMLLVMA-FIATDIFI 77
QY 109 SONTQDDLRKYVFLNKMVGRGLSICPCLLSVLQAI-SPSIFSLAKLKHPSASHILGF 167
DB 78 SWRGWDDIICKFLVLYRVLRGLSLCTTSMLSVLQAILSPRSCLAKFKRKSLLHISCA 137
QY 168 FLFSWLNMFIVGFCCTLRPPVKRGQSSVCHTALFLFAHELHPQETVFTNDF---E 223
DB 138 ILFSLVLYMLIGSQLSVIIATP-----DYFIOTI-----RGYLSAF-----TO-----PACPRV 172
QY 224 GCHLYRVHGPLKRLHG-----DYFIOTI-----RGYLSAF-----TO-----PACPRV 261
DB 173 SCSTLPUSYVQSMFSTLLVIRDVFLISLWVSTWYVALLCRRHKKTQHLQGISLSPKT 232
QY 262 SPVKRASQAILLVSVFVFTVWDFTFSGGVTVWINDSLVWLQVIVANSYAAISPLMLI 321
DB 233 SPKQATQTLMLMSFFVLMTIYDVIVSCSRTMFLNDPTSYNMQIFVYVHIYATVSPFVM 292
QY 322 YADNQIEKTLQ 332
DB 293 STEKHIVNCLR 303

RESULT 11
Q9EQ49 Q9EQ49 PRELIMINARY; PRT; 313 AA.
AC Q9EQ49;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VOMERONASAL RECEPTOR VIRAG.
GN VIRAG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=129/SVJ;
RX MEDLINE=20568485; PubMed=11116090;
RA Del Punta K., Rothman A., Rodriguez I., Mombaerts P.;
RT "Sequence diversity and genomic organization of vomeronasal receptor
RT genes in the mouse.";
RL Genome Res. 10:1958-1967(2000).
DR EMBL; AF291485; AAC42079.1; -.
DR MGD; MGI:2148511; V1ra6.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 313 AA; 35584 MW; 04EDF79B5FADB17C CRC64;

Query Match 17.8%; Score 330; DB 11; Length 313;
Best Local Similarity 29.4%; Pred. No. 3.6e-22;
Matches 99; Conservative 47; Mismatches 113; Indels 78; Gaps 9;

QY 49 FYQASFGISANTILLFHIFTVFHSRKSIDMIISHLSLHILLFTQAILVSLDFFG 108
DB 17 FFSEIGIGISGNSFLFLHILKFIHGHRPRLTDLPIGLLSLHLLMLLV-AAFIATDIFI 75
QY 109 SONTQDDLRKYVFLNKMVGRGLSICPCLLSVLQAI-SPSIFSLAKLKHPSASHILGF 167

```

```
Db 76 SRGWDDIICKELVLYRVLGRFSLCTTSMLSILOAILSPRSSCLAKFKHISPHHISGA 135
Qy 168 FLFSVNLNMFVIGVFCCTLRLLPPVKGQSSVCHTALFLFAHELHPQETVETVHNDF---E 223
Db 136 ILFLSVLYMLIG-----SOLLVSIATPNTLM-----NDFIYVTO 170
Qy 224 GCHLYRVHGPKRLKRGDYFIQIRGYLSAFTQ-----255
Db 171 SCSIL---PL-----SVLMQSIYSTLLAIREFFLISLWLSNWMYVALLSMHRKQTOHL 221
Qy 256 ---PACPRVSPVKRASQAILLVSVFTYVWDFTSFSGGVTVWINDSLVWLVQIVANSY 312
Db 222 HGTNLSPKSPKSPQSAQTQTLMLISFLLMTIYDVTSCSRMTFLNDPTSYSIELFIMHIY 281
Qy 313 AAISPLMLIYADNQIFKTLQMLWFKYLSPPKMLKFN 349
Db 282 ATVSPFVFMSTGKHIVNFKSM-----GKRVINFN 311
```

```
RESULT 12
Q62852
ID Q62852 PRELIMINARY; PRT; 311 AA.
AC Q62852;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PHEROMONE RECEPTOR VN3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=VOMERONASAL ORGAN;
RX MEDLINE=96028094; PubMed=7585937;
RA Dutac C., Axel R.;
RT "A novel family of genes encoding putative pheromone receptors in
  mammals.";
RL Cell 83:195-206(1995).
DR EMBL; U36895; AAC52284.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR PRINTS; PR01534; VOMERONASLIR.
DR PROSITE; PS02362; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 311 AA; 35426 MW; F5997D351B33F318 CRC64;
```

```
Query Match 17.8%; Score 329.5; DB 11; Length 311;
Best Local Similarity 29.5%; Pred. No. 4e-22;
Matches 95; Conservative 48; Mismatches 118; Indels 61; Gaps 8;

Qy 44 TIQIFYPQASGISANTILLFLHFIFFVFSHRKSIDMIISHLISLIHLLFTQAILVS 103
Db 13 TIRNFTSEIGILANSFLLLHFIFKFIKRGORSRLTDLPGICLLSLIHLMLMGA-FIA 71
Qy 104 LDFFGSQNTQDRLRYKVIYFLNKNVRLGSLCTPCLLSVLOAI-ISPSIFSLAKLHPSAS 162
Db 72 IDIFTSWRGWDDICKFLVLYRSFRLGSLCTCTMLSVLQAI-TLSPRSSCLAKFKHSPH 131
Qy 163 HILGFFLFSWLNMFVIGVFCCTLRLLPPVKGQSSVCHTALFLFAHELHPQETVETVHNDF 222
Db 132 HVSCAISLSILYMEISSHLVSINATP-----NLTNNF 166
Qy 223 ----ECCHLYRVHGPKRLKRGDYF--IQTIR--GYLSAFTQAC-----258
Db 167 MQVTOscyII-----PLSYLMQSMFSTLLAIRDISILSVLSTCYNEVLLCHRRNOIQL 222
Qy 259 -----PRVSPVKRASQAILLVSVFTYVWDFTSFSGGVTVWINDSLVWLVQIVANSY 312
Db 223 QGYNLSPKASPRQATQTLMLTMTFFVLMSIFDSIYSCSRMTYLNPTSYSYIQIFVVDIY 282
```

```
Qy 313 AAISPLMLIYADNQIFKTLQML 334
Db 283 ATVSPFVFMSTGKHIVNFKSM 304

RESULT 13
Q9EQ46
ID Q9EQ46 PRELIMINARY; PRT; 310 AA.
AC Q9EQ46;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE VOMERONASAL RECEPTOR VIRB7.
GN VIRB7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20568485; PubMed=11116090;
RA Del Punta K., Rothman A., Rodriguez I., Mombaerts P.;
RT "Sequence diversity and genomic organization of vomeronasal receptor
  genes in the mouse.";
RL Genome Res. 10:1958-1967(2000).
DR EMBL; AF291493; AAG42087.1; -.
DR MGI; MGI:2148518; Virb7.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 310 AA; 35206 MW; D60E8D958E961435 CRC64;
```

```
Query Match 17.4%; Score 322.5; DB 11; Length 310;
Best Local Similarity 28.3%; Pred. No. 1.7e-21;
Matches 90; Conservative 56; Mismatches 107; Indels 65; Gaps 9;

Qy 45 IQIFYPQASGISANTILLFLHFIFFVFSHRKSIDMIISHLISLIHLLFTQAILVS 104
Db 13 MQIILFSEVSGISANSILFIAHVCMLGPNRPKPIDLYIAFLSUTQLMLLTMG-LIAV 71
Qy 105 DFFGSQNTQDRLRYKVIYFLNKNVRLGSLCTPCLLSVLOAI-ISPSIFSLAKLHPSASH 163
Db 72 DMFLSQGIWSDTTCOSLIYHLRLGLSLCATCLLNILMTITLSSRSFCSTFKFKHSPHH 131
Qy 164 ILGFFLFSWLN-----NMFVIGVFCCTLRLLPPVKGQSSVCHTALFLFAHELHPQETV 217
Db 132 ISGAFIFCVLYMSFSSHLFISII-----ATHNLTSENFY 167
Qy 218 HTNDFEGCHLYRVHGPKRLKRGDYFI--QTIR-----GYLSAF-----253
Db 168 VT---QSCSL---PLSVSRTSMFSPAPMAIREAFVLSLMSLSSGYMVALLWRHKKQAQ 220
Qy 254 --TQCAPRVSPVKRASQAILLVSVFTYVWDFTSFSGGVTVWINDSLVWLVQIVANS 311
Db 221 LHSTSLSKASPEQRATRTILLMSFFVLYILENAVFYSRIKFKDGSILCYVQIILCHS 280
Qy 312 YAAISPLMLIYADNQIEK 329
Db 281 YATVNPFFVICTEKHIK 298

RESULT 14
Q9EP79
ID Q9EP79 PRELIMINARY; PRT; 309 AA.
AC Q9EP79;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE VN3 (VOMERONASAL RECEPTOR VIRA7).
GN VIRA7.
OS Mus musculus (Mouse).
```


